



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : C12N 15/62, 15/13, 15/14 C12N 15/12, C07K 13/00 A61K 37/02, C12N 5/10	A2	(11) International Publication Number: WO 91/08298 (43) International Publication Date: 13 June 1991 (13.06.91)
(21) International Application Number: PCT/US90/06849 (22) International Filing Date: 21 November 1990 (21.11.90) (30) Priority data: 444,625 22 November 1989 (22.11.89) US (71) Applicant: GENENTECH, INC. [US/US]; 460 Point San Bruno Boulevard, South San Francisco, CA 94080 (US). (72) Inventors: CAPON, Daniel, J. ; 817 Oregon Street, San Mateo, CA 94402 (US). LASKY, Laurence, A. ; Star Route Box 460, Sausalito, CA 94965 (US). (74) Agents: ADLER, Carolyn, R. et al.; Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080 (US).		(81) Designated States: AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: HYBRID IMMUNOGLOBULINS (57) Abstract Novel polypeptides are provided, together with methods for making and using them, and nucleic acids encoding them. These polypeptides are useful as cell surface adhesion molecules and ligands, and are useful in therapeutic or diagnostic compositions and methods		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FI	Finland	ML	Mali
AU	Australia	FR	France	MN	Mongolia
BB	Barbados	GA	Gabon	MR	Mauritania
BE	Belgium	GB	United Kingdom	MW	Malawi
BF	Burkina Faso	GN	Guinea	NL	Netherlands
BG	Bulgaria	GR	Greece	NO	Norway
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	SD	Sudan
CF	Central African Republic	KP	Democratic People's Republic of Korea	SE	Sweden
CG	Congo	KR	Republic of Korea	SN	Senegal
CH	Switzerland	LI	Liechtenstein	SU	Soviet Union
CI	Côte d'Ivoire	LK	Sri Lanka	TD	Chad
CM	Cameroon	LU	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	US	United States of America
DK	Denmark	MG	Madagascar		
ES	Spain				

HYBRID IMMUNOGLOBULINS

Background of the Invention

This invention relates to novel ligand binding molecules and receptors, and to compositions and methods for improving the circulating plasma half-life of ligand binding molecules. In particular, this invention also relates to hybrid immunoglobulin molecules, to methods for making and using these immunoglobulins, and to nucleic acids encoding them.

Immunoglobulins are molecules containing polypeptide chains held together by disulfide bonds, typically having two light chains and two heavy chains. In each chain, one domain (V) has a variable amino acid sequence depending on the antibody specificity of the molecule. The other domains (C) have a rather constant sequence common among molecules of the same class. The domains are numbered in sequence from the amino-terminal end.

The immunoglobulin gene superfamily consists of molecules with immunoglobulin-like domains. Members of this family include class I and class II major histocompatibility antigens, immunoglobulins, T-cell receptor α , β , γ and δ chains, CD1, CD2, CD4, CD8, CD28, the γ , δ and ϵ chains of CD3, OX-2, Thy-1, the intercellular or neural cell adhesion molecules (I-CAM or N-CAM), lymphocyte function associated antigen-3 (LFA-3), neurocytoplasmic protein (NCP-3), poly-Ig receptor, myelin-associated glycoprotein (MAG), high affinity IgE receptor, the major glycoprotein of peripheral myelin (Po), platelet derived growth factor receptor, colony stimulating factor-1 receptor, macrophage Fc receptor, Fc gamma receptors and carcinoembryonic antigen.

It is known that one can substitute variable domains (including hypervariable regions) of one immunoglobulin for another, and from one species to another. See, for example, EP O 173 494; EP O 125 023; Munro, *Nature* 312: (13 December 1984); Neuberger *et al.*, *Nature* 312: (13 December 1984); Sharon *et al.*, *Nature* 309: (24 May 1984); Morrison *et al.*, *Proc. Natl. Acad. Sci. USA* 81:6851-6855 (1984); Morrison *et al.*, *Science* 229:1202-1207 (1985); and Boulianne *et al.*, *Nature* 312:643-646 (13 December 1984).

Morrison *et al.*, *Science* 229:1202-1207 (1985) teaches the preparation of an immunoglobulin chimera having a variable region from one species fused to an immunoglobulin constant region from another species. This reference suggests molecules having immunoglobulin sequences fused with non-immunoglobulin sequences (for example enzyme sequences), however the reference teaches only immunoglobulin variable domains attached to the non-immunoglobulin sequence. Morrison *et al.*, EP O 173 494 teaches a similar chimera. While the term "receptor" is used by the authors, and the background section refers to "receptors such as immunoglobulins, enzymes and membrane proteins", the stated "receptors of interest" include "B-cell and T-cell receptors, more particularly, immunoglobulins, such as IgM, IgG, IgA, IgD and IgE, as well as the various subtypes of the individual groups" (page 3 lines 10-13). The disclosure of this reference is specific to immunoglobulin chimeras (see for example page 3, lines 21-30).

It has also been shown that it is possible to substitute immunoglobulin variable-like domains from two members of the immunoglobulin gene superfamily--CD4 and the T cell receptor--for a variable domain in an immunoglobulin; see *e.g.* Capon *et al.*, Nature 337:525-531, 1989, Traunecker *et al.*, Nature 339:68-70, 1989, Gascoigne *et al.*, Proc. Nat. Acad. Sci. 84:2936-2940, 1987, and published European application EPO 0 325 224 A2.

A large number of proteinaceous substances are known to function by binding specifically to target molecules. These target molecules are generally, but need not be, proteins. The substances which bind to target molecules or ligands are referred to herein as ligand binding partners, and include receptors and carrier proteins, as well as hormones, cellular adhesive proteins, tissue-specific adhesion factors, lectin binding molecules, growth factors, enzymes, nutrient substances and the like.

Lymphocytes are examples of cells which are targeted to specific tissues. Lymphocytes are mediators of normal tissue inflammation as well as pathologic tissue damage such as occurs in rheumatoid arthritis and other autoimmune diseases. Vertebrates have evolved a mechanism for distributing lymphocytes with diverse antigenic specificities to spatially distinct regions of the organism (Butcher, E. C., Curr. Top. Micro. Immunol. 128, 85 (1986); Gallatin, W. M., *et al.*, Cell 44, 673 (1986); Woodruff, J. J., *et al.*, Ann. Rev. Immunol. 5, 201 (1987); Duijvestijn, A., *et al.*, Immunol. Today 10, 23 (1989); Yednock, T. A., *et al.*, Adv. Immunol. (in press) (1989)).

This mechanism involves the continuous recirculation of the lymphocytes between the blood and the lymphoid organs. The migration of lymphocytes between the blood, where the cells have the greatest degree of mobility, and the lymphoid organs, where the lymphocytes encounter sequestered and processed antigen, is initiated by an adhesive interaction between receptors on the surface of the lymphocytes and ligands on the endothelial cells of specialized postcapillary venules, *e.g.*, high endothelial venules (HEV) and the HEV-like vessels induced in chronically inflamed synovium.

The lymphocyte adhesion molecules have been specifically termed homing receptors, since they allow these cells to localize in or "home" to particular secondary lymphoid organs.

Candidates for the lymphocyte homing receptor have been identified in mouse, rat and human (Gallatin, W. M., *et al.*, Nature 303, 30 (1983) Rasmussen, R. A., *et al.*, J. Immunol. 135, 19 (1985); Chin, Y. H., *et al.*, J. Immunol. 136, 2556 (1986); Jalkanen, S., *et al.*, Eur. J. Immunol. 10, 1195 (1986)). The following literature describes work which has been done in this area through the use of a monoclonal antibody, termed Mel 14, directed against a purported murine form of a lymphocyte surface protein (Gallatin, W. M., *et al.*, *supra*; (Mountz, J. D., *et al.*, J. Immunol. 140, 2943 (1988); (Lewinsohn, D. M., *et al.*, J. Immunol. 138, 4313 (1987); Siegelman, M., *et al.*, Science 231, 823 (1986); St. John, T., *et al.*, Science 231, 845 (1986)).

Immunoprecipitation experiments have shown that this antibody recognizes a diffuse, ~90,000 dalton cell surface protein on lymphocytes (Gallatin, W. M., *et al.*, *supra*) and a ~100,000 dalton protein on neutrophils (Lewinsohn, D. M., *et al.*, *supra*).

5 A partial sequence--13 residues--for a purported lymphocyte homing receptor identified by radioactively labeled amino acid sequencing of a Mel-14 antibody-defined glycoprotein was disclosed by Siegelman *et al.* (Siegelman, M., *et al.*, Science 231, 823 (1986)).

Lectins are proteins with a carbohydrate-binding domain found in a variety of animals, including humans as well as the acorn barnacle and the flesh fly. The concept of lectins functioning in cell adhesion is exemplified by the interaction of certain viruses and bacteria with eucaryotic host cells (Paulson, J. C., The Receptors Vol. 2 P. M. Conn, Eds. (Academic Press, NY, 1985), pp. 131; Sharon, N., FEBS Lett. 217, 145 (1987)). In eucaryotic cell-cell interactions, adhesive functions have been inferred for endogenous lectins in a variety of systems (Grabel, L., *et al.*, Cell 17, 477 (1979); Fenderson, B., *et al.*, J. Exp. Med. 160, 1591 (1984); Kunemund, V., J. Cell Biol. 106, 213 (1988); Bischoff, R., J. Cell Biol. 102, 2273 (1986); Crocker, P. R., *et al.*, J. Exp. Med. 164, 1862 (1986); including invertebrate (Glabe, C. G., *et al.*, J. Cell Biol. 94, 123 (1982); DeAngelis, P., *et al.*, J. Biol. Chem. 262, 13946 (1987)) and vertebrate fertilization (Bleil, J. D., *et al.*, Proc. Natl. Acad. Sci., U.S.A. 85, 6778 (1988); Lopez, L. C., *et al.*, J. Cell Biol. 101, 1501 (1985)). The use of protein-sugar interactions as a means of achieving specific cell recognition appears to be well known.

20 The literature suggests that a lectin may be involved in the adhesive interaction between the lymphocytes and their ligands (Rosen, S. D., *et al.*, Science 228, 1005 (1985); Rosen, S. D., *et al.*, J. Immunol. (in press) (1989); Stoolman, L. M., *et al.*, J. Cell Biol. 96, 722 (1983); Stoolman, L. M., *et al.*, J. Cell Biol. 99, 1535 (1984); Yednock, T. A., *et al.*, J. Cell Bio. 104, 725 (1987); Stoolman, L. M., *et al.*, Blood 70, 1842 (1987); A related approach by Brandley, B. K., *et al.*, J. Cell Biol. 105, 991 (1987); Yednock, T. A., *et al.*, in preparation; and Yednock, T. A., *et al.*, J. Cell Biol. 104, 725 (1987)).

30 The character of a surface glycoprotein that may be involved in human lymphocyte homing was investigated with a series of monoclonal and polyclonal antibodies generically termed Hermes. These antibodies recognized a ~90,000 dalton surface glycoprotein that was found on a large number of both immune and non-immune cell types and which, by antibody pre-clearing experiments, appeared to be related to the Mel 14 antigen. (Jalkanen, S., *et al.* Ann. Rev. Med., 38, 467-476 (1987); Jalkanen, S., *et al.*, Blood, 66 (3), 577-582 (1985); Jalkanen, S., *et al.*, J. Cell Biol., 105, 983-990 (1987); Jalkanen, S., *et al.*, Eur. J. Immunol., 18, 1195-1202 (1986)).

35 Epidermal growth factor-like domains have been found on a wide range of proteins, including growth factors, cell surface receptors, developmental gene products, extracellular matrix proteins, blood clotting factors, plasminogen activators, and complement (Doolittle, R. F., *et al.*, CSH Symp. 51, 447 (1986)).

-4-

A lymphocyte cell surface glycoprotein (referred to hereafter as the "LHR") has been characterized which mediates the binding of lymphocytes to the endothelium of lymphoid tissue. Full length cDNA clones and DNA encoding the human and the murine LHR (HuLHR and MLHR, respectively) have been identified and isolated, and moreover this DNA is readily expressed by recombinant host cells. The nucleotide and amino acid sequence of the human LHR (HuLHR) is shown in Fig. 1. The nucleotide and amino acid sequence of the murine LHR (MLHR) is shown in Fig. 2. Also provided are LHR having variant amino acid sequences or glycosylation not otherwise found in nature, as well as other derivatives of the LHR having improved properties including enhanced specific activity and modified plasma half-life, as well as enabling methods for the preparation of such variants.

It is shown herein that the LHR is a glycoprotein which contains the following protein domains: a signal sequence, a carbohydrate binding domain, an epidermal growth factor-like (egf) domain, at least one and preferably two complement binding domain repeat, a transmembrane binding domain (TMD), and a charged intracellular or cytoplasmic domain. The LHR of this invention contains at least one but not necessarily all of these domains.

A successful strategy in the development of drugs for the treatment of many abnormalities in ligand-binding partner interactions has been the identification of antagonists which block the binding or interaction between ligand and binding partner. One approach has been to use an exogenous binding partner as a competitive antagonist for the native binding partner. However, many ligand binding partners are cell membrane proteins which are anchored in the lipid bilayer of cells. The presence of membrane components is typically undesirable from the standpoint of manufacturing and purification. In addition, since these molecules are normally present only on cell surfaces, it would be desirable to produce them in a form which is more stable in the circulation. Additionally, even truncated or soluble ligand binding partners may not be optimally effective as therapeutics since they possess a relatively short *in vivo* plasma half-life, may not cross the placental or other biological barriers, and since merely sequestering their ligand recognition site without delivering an effector function may be inadequate for therapeutic purposes.

Accordingly, it is an object of this invention to produce ligand binding partners fused to moieties which serve to prolong the *in vivo* plasma half-life of the ligand binding partner, such as immunoglobulin domains or plasma proteins, and facilitate its purification by protein A. It is a further object to provide novel hybrid immunoglobulin molecules which combine the adhesive and targeting characteristics of a ligand binding partner with immunoglobulin effector functions such as complement binding, cell receptor binding and the like. Yet another object is to provide molecules with novel functionalities such as those described above for therapeutic use, or for use as diagnostic reagents for the *in vitro* assay of the ligand binding partners or their targets. It is another object to provide multifunctional molecules in which a plurality of ligand binding partners (each of which may be the same or different) are

-5-

assembled, whereby the molecules become capable of binding and/or activating more than one ligand.

In particular, it is an objective to prepare molecules for directing ligand binding partners such as toxins, cell surface partners, enzymes, nutrient substances, growth factors, hormones or effector molecules such as the constant domain-like portions of a member of the immunoglobulin gene superfamily to cells bearing ligands for the ligand binding partners, and for use in facilitating purification of the ligand binding partners.

Another object of this invention is to provide ligand binding partner-immunoglobulin hybrid heteropolymers, especially heterodimers and heterotetramers, which are used in the targeting of therapeutic moieties to specific tissues and ligands. For example, a hybrid immunoglobulin consisting of one LHR-IgG chain and one CD4-IgG chain can be used to target CD4-IgG to tissues infected by viruses such as the human immunodeficiency virus (HIV). Similarly, a molecule having a ligand binding partner-plasma protein portion combined with a toxin-plasma protein portion is used to deliver the toxin to desired tissues.

It is another object to provide a method for expression of these molecules in recombinant cell culture.

Summary of the Invention

The objects of this invention are accomplished by providing novel polypeptides comprising a ligand binding partner fused to a stable plasma protein which is capable of extending the *in vivo* plasma half-life of the ligand binding partner when present as a fusion with the ligand binding partner, in particular wherein such a stable plasma protein is an immunoglobulin constant domain. DNA encoding the polypeptides, cultures and methods for making the polypeptides are also provided.

In most cases where the stable plasma protein is normally found in a multimeric form, e.g., immunoglobulins or lipoproteins, in which the same or different polypeptide chains are normally disulfide and/or noncovalently bound to form an assembled multichain polypeptide, the fusions herein containing the ligand binding partner also will be produced and employed as a multimer having substantially the same structure as the stable plasma protein precursor. These multimers will be homogeneous with respect to the ligand binding partner they comprise, or they may contain more than one ligand binding partner. Furthermore, they may contain one or more ligand binding partner moieties.

In a preferred embodiment in which the stable plasma protein is an immunoglobulin chain, the ligand binding partner will be substituted into at least one chain, and ordinarily for the variable region of the immunoglobulin or suitable fragment thereof. However, it will be understood that this invention also comprises those fusions where the same or different ligand binding partners are substituted into more than one chain of the immunoglobulin. If the ligand binding partners are different, then the final assembled multichain polypeptide is

-6-

capable of crosslinking ligands in a fashion that may not be possible with multifunctional antibodies having native variable regions.

A particular multichain fusion of this sort is one in which the variable region of one immunoglobulin chain has been substituted by the ligand binding region of a first receptor such as CD4 while the variable region of another immunoglobulin chain has been substituted by a binding functionality of the LHR, both immunoglobulin chains being associated with one another in substantially normal fashion.

The fusions of this invention may be further modified by linking them through peptidyl or *in vitro* generated bonds to an additional therapeutic moiety such as a polypeptide toxin, a diagnostic label or other functionality.

The fusions of this invention are made by transforming host cells with nucleic acid encoding the fusion, culturing the host cell and recovering the fusion from the culture. Also provided are vectors and nucleic acid encoding the fusion, as well as therapeutic and diagnostic compositions comprising them.

In certain respects this invention is directed to LHR *per se*. The LHR of this invention is full-length, mature LHR, having the amino acid sequence described herein at Figs. 1 and 2, and naturally occurring alleles, covalent derivatives made by *in vitro* derivatization, or predetermined amino acid sequence or glycosylation variants thereof.

The novel compositions provided herein are purified and formulated in pharmacologically acceptable vehicles for administration to patients in need of antiviral, neuromodulatory or immunomodulatory therapy, and for use in the modulation of cell adhesion. This invention is particularly useful for the treatment of patients having receptor-mediated abnormalities. In addition, the compositions provided herein are useful intermediates in the purification of the ligand binding partner from recombinant cell culture, wherein antibodies or other substances capable of binding the stable plasma protein component are used to absorb the fusion, or are useful in diagnostic assays for the ligand binding partner wherein the stable plasma protein serves as an indirect label.

Brief Description of the Figures

Figure 1 depicts the amino acid and DNA sequence of the Human LHR (HuLHR).

Figure 2 depicts the amino acid and DNA sequence of the Murine LHR (MLHR).

Figure 3 shows a comparison between the amino acid sequences for the mature HuLHR and MLHR.

Figures 4A-4B show the results of isolation and N-terminal sequencing of the MLHR. Fig. 4A shows the gas phase Edman degradation results of a 90,000 dalton band from an SDS-polyacrylamide gel of material purified from a detergent extract of murine spleens by Mel 14 monoclonal antibody affinity chromatography. The residues underlined between amino acids 7 and 15 were chosen to produce the oligonucleotide probe shown in Fig. 4B. Fig. 4B shows as 32-fold redundant 26-mer oligonucleotide probe.

-7-

Figure 5 shows the transient expression of the MLHR cDNA clone. Lanes A - F signify the following: --A. Lysates of 293 cells transfected with a MLHR expression plasmid immunoprecipitated with Mel 14 monoclonal antibody. --B. Supernatants of 293 cells transfected with a MLHR expression plasmid immunoprecipitated with Mel 14 monoclonal antibody. --C. Lysates of 293 cells transfected with a plasmid expressing the HIV gp120 envelope glycoprotein immunoprecipitated with the Mel 14 monoclonal antibody. --D. Supernatants of 293 cells transfected with the HIV envelope expression plasmid immunoprecipitated with the Mel 14 monoclonal antibody. --E. Supernatants of 38C13 cells immunoprecipitated with the Mel 14 monoclonal antibody. --F. Lysates of 38C13 cells surface labeled with I¹²⁵ and immunoprecipitated with the Mel 14 monoclonal antibody.

Figures 6A-6C show protein sequences which are heterologous but functionally comparable to the MLHR. Those lines labelled "MLHR" correspond to the MLHR of Fig. 2. Fig. 6A compares carbohydrate-binding domains; Fig. 6B compares epidermal growth factor domains; and Fig. 6C compares complement binding factor domains.

Figure 7 is a schematic of protein domains found in the LHR, including the signal sequence, carbohydrate binding domain, epidermal growth factor (egf) domain, two complement binding domain repeats (arrows), transmembrane binding domain (TMD), and charged intracellular domain.

Figure 8 shows the construction of MLHR-IgG chimeras containing the lectin, lectin-egf, and lectin-egf-complement regulatory motifs.

Figure 9 shows the gel electrophoresis of the products of the expression and purification of the MLHR-IgG chimeras.

Figure 10 shows polyphosphomannan ester (PPME) binding analysis of various MLHR-IgG chimeras.

25 Detailed Description

Ligand binding partners as defined herein are proteins known to function to bind specifically to target ligand molecules, and are generally found in their native state as secreted or membrane bound polypeptides; membrane-bound ligand binding partners typically include a hydrophobic transmembrane region or phospholipid anchor. Ligand binding partners include receptors and carrier proteins, as well as hormones, cellular adhesive proteins (proteins which direct or induce the adhesion of one cell to another), lectin binding molecules, growth factors, enzymes, nutrient substances and the like. CD antigens which are not members of the immunoglobulin gene superfamily or otherwise excluded as set forth above are suitable ligand binding partners, Knapp *et al.* Immunology Today 10 (8):253-258, 1989. The platelet growth factor receptor and insulin receptor may optionally be ligand binding partners. Ligand binding partners include not only the full length native form, but truncations or other amino acid sequence variants that remain capable of binding to the normal ligand.

As used herein, the term "ligand binding partner" specifically excludes polymorphic and

nonpolymorphic members of the immunoglobulin gene superfamily, and proteins which are homologous thereto, such as class I and class II major histocompatibility antigens, immunoglobulins, T-cell receptor α , β , γ and δ chains, CD1, CD2, CD4, CD8, CD28, the γ , δ and ϵ chains of CD3, OX-2, Thy-1, the intercellular or neural cell adhesion molecules (I-CAM or N-CAM), lymphocyte function associated antigen-3 (LFA-3), neurocytoplasmic protein (NCP-3), poly-Ig receptor, myelin-associated glycoprotein (MAG), high affinity IgE receptor, the major glycoprotein of peripheral myelin (Po), platelet derived growth factor receptor, colony stimulating factor-1 receptor, macrophage Fc receptor, Fc gamma receptors and carcinoembryonic antigen. Homologous to a member of the immunoglobulin gene superfamily, for the purposes of this exclusion only, means having the sequence of a member of the immunoglobulin gene superfamily or having a sequence therewithin which has substantially the same (or a greater degree of) amino acid sequence homology to a known member of the superfamily as the specific examples given above have to the sequence of an immunoglobulin variable or constant domain. Note that this does not exclude embodiments in which a ligand binding partner fusion is assembled into a multimer with, in addition, a member or fusion of a member of the immunoglobulin gene superfamily.

Also specifically excluded from the term "ligand binding partner" are multiple subunit (chain) polypeptides encoded by discrete genes (genes which do not encode a single chain precursor polypeptide leading to the multiple subunit polypeptide), with at least one subunit of the polypeptide being ordinarily inserted into the cell membrane, including cellular receptors (e.g., integrins) for extracellular matrix molecules, as exemplified in PCT Pub. WO 90/06953 published June 28, 1990. Note that this does not exclude embodiments in which a ligand binding partner fusion is assembled into a multimer with, in addition, a multiple subunit polypeptide or fusion of a multiple subunit polypeptide as defined in this paragraph.

Stable plasma proteins are proteins typically having about from 30 to 2,000 residues, which exhibit in their native environment an extended half-life in the circulation, *i.e.* greater than about 20 hours. Examples of suitable stable plasma proteins are immunoglobulins, albumin, lipoproteins, apolipoproteins and transferrin. The ligand binding partner typically is fused to the plasma protein at the N-terminus of the plasma protein or fragment thereof which is capable of conferring an extended half-life upon the ligand binding partner. The ligand binding partner generally is fused at its native C-terminus to the plasma protein. However, on occasion it may be advantageous to fuse a truncated form of the ligand binding partner (in which the transmembrane and cytoplasmic regions have been deleted) to a portion of the stable protein that exhibits a substantially hydrophobic hydropathy profile, typically the first site in the mature stable protein in which a hydrophobic region having greater than about 20 residues appears. Such sites are present in transferrin and are quite common in albumin and apolipoproteins and should present no difficulty in identification. As much of the remainder of the stable protein as is required to confer extended plasma half-life on the ligand

binding partner is then incorporated into the fusion. Increases of greater than about 100 % on the plasma half-life of the ligand binding partner are satisfactory.

In some preferred embodiments, the binding partner is an LHR. The LHR is defined as a polypeptide having a qualitative biological activity in common with the LHR of Fig. 1 or Fig. 2, and which preferably contains a domain greater than about 70% homologous with the carbohydrate binding domain, the epidermal growth factor domain, or the carbohydrate binding domain of the LHR of Fig. 1 or Fig. 2.

Homology with respect to a LHR is defined herein as the percentage of residues in the candidate sequence that are identical with the residues in the carbohydrate binding domain, the epidermal growth factor domain, or the complement binding domains in Fig. 1 or Fig. 2 after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent homology.

Included within the scope of the LHR as that term is used herein are LHRs having the amino acid sequences of the HuLHR or MLHR as set forth in Fig. 1 or 2, deglycosylated or unglycosylated derivatives of the LHR, homologous amino acid sequence variants of the sequence of Fig. 1 or 2, and homologous *in vitro*-generated variants and derivatives of the LHR, which are capable of exhibiting a biological activity in common with the LHR of Fig. 1 or Fig. 2.

LHR or LHR-fragment biological activity is defined as either 1) immunological cross-reactivity with at least one epitope of the LHR, or 2) the possession of at least one adhesive, regulatory or effector function qualitatively in common with the LHR.

One example of the qualitative biological activities of the LHR is its binding to ligands on the specialized high endothelial cells of the lymphoid tissues. Also, it frequently requires a divalent cation such as calcium for ligand binding.

Immunologically cross-reactive as used herein means that the candidate polypeptide is capable of competitively inhibiting the qualitative biological activity of the LHR having this activity with polyclonal antisera raised against the known active analogue. Such antisera are prepared in conventional fashion by injecting goats or rabbits, for example, subcutaneously with the known active analogue in complete Freund's adjuvant, followed by booster intraperitoneal or subcutaneous injection in incomplete Freund's.

Structurally, as shown in Figure 3, the LHR includes several domains which are identified as follows (within ± 10 residues): a signal sequence (residues 20-32), which is followed by a carbohydrate binding domain (identified in Fig. 3 as a "lectin" domain) (residues 39-155), an epidermal growth factor (egf) domain (residues 160-193), a complement factor binding domain (residues 197-317), a transmembrane binding domain (TMD) (residues 333-355), and a cytoplasmic domain (residues 356-372). The boundary for the LHR extracellular domain generally is at, or within about 30 residues of, the N-terminus of the transmembrane

domain, and is readily identified from an inspection of the LHR sequence. Any or all of these domains are utilized in the practice of this invention.

Figs. 6A - 6C show a variety of proteins having some homology to three of these domains. Fig. 6A shows carbohydrate binding domains, Fig. 6B shows epidermal growth factor domains, and Fig. 6C shows somewhat homologous complement binding domains.

A comparison of the amino sequences of HuLHR and MLHR is presented in Fig. 3, and shows a high degree of overall sequence homology (~83%). The degrees of homology between the various domains found in the HuLHR versus the MLHR, however, are variable. For example, the degree of sequence conservation between the MLHR and the HuLHR in both the carbohydrate-binding and egf domains is approximately 83%, while the degree of conservation in the first complement binding repeat falls to 79% and only 63% in the second repeat, for an overall complement binding domain homology of ~71%. Furthermore, while the two MLHR complement binding domain repeats are identical, those in the HuLHR have differences, and differ as well to the murine repeats. Interestingly, the degree of conservation between the two receptors in the transmembrane sequence and surrounding regions is virtually identical, with only one conservative hydrophobic substitution, probably within the transmembrane anchor region.

The surface glycoprotein discussed above that is recognized by the series of monoclonal and polyclonal antibodies generically termed Hermes is specifically excluded from the scope of this invention.

Immunoglobulins and certain variants thereof are known and many have been prepared in recombinant cell culture. For example, see U.S. Patent 4,745,055; EP 256,654; Faulkner *et al.*, Nature 298:286 (1982); EP 120,694; EP 125,023; Morrison, J. Immun. 123:793 (1979); Köhler *et al.*, P.N.A.S. USA 77:2197 (1980); Raso *et al.*, Cancer Res. 41:2073 (1981); Morrison *et al.*, Ann. Rev. Immunol. 2:239 (1984); Morrison, Science 229:1202 (1985); Morrison *et al.*, P.N.A.S. USA 81:6851 (1984); EP 255,694; EP 266,663; and WO 88/03559. Reassorted immunoglobulin chains also are known. See for example U.S. patent 4,444,878; WO 88/03565; and EP 68,763 and references cited therein.

Ordinarily, the ligand binding partner is fused C-terminally to the N-terminus of the constant region of immunoglobulins in place of the variable region(s) thereof, however N-terminal fusions of the binding partner are also desirable. The transmembrane regions or lipid or phospholipid anchor recognition sequences of ligand binding partners comprising such regions or sequences are preferably inactivated or deleted prior to fusion.

Typically, such fusions retain at least functionally active hinge, CH2 and CH3 domains of the constant region of an immunoglobulin heavy chain. Fusions are also made to the C-terminus of the Fc portion of a constant domain, or immediately N-terminal to the CH1 of the heavy chain or the corresponding region of the light chain. This ordinarily is accomplished by constructing the appropriate DNA sequence and expressing it in recombinant cell culture.

Alternatively, however the polypeptides of this invention may be synthesized according to known methods.

The precise site at which the fusion is made is not critical; particular sites are well known and may be selected in order to optimize the biological activity, secretion or binding characteristics of the binding partner. The optimal site will be determined by routine experimentation.

In some embodiments the hybrid immunoglobulins are assembled as monomers, or hetero- or homo-multimers, and particularly as dimers or tetramers. Generally, these assembled immunoglobulins will have known unit structures as represented by the following diagrams. A basic four chain structural unit is the form in which IgG, IgD, and IgE exist. A four chain unit is repeated in the higher molecular weight immunoglobulins; IgM generally exists as a pentamer of basic four-chain units held together by disulfide bonds. IgA globulin, and occasionally IgG globulin, may also exist in a multimeric form in serum. In the case of multimers, each four chain unit may be the same or different.

In the diagrams herein, "A" means at least a portion of a ligand binding partner containing a ligand binding site which is capable of binding its ligand; X is an additional agent, which may be another functional ligand binding partner (same as A or different), a multiple subunit (chain) polypeptide as defined above (e.g., an integrin), a portion of an immunoglobulin superfamily member such as a variable region or a variable region-like domain, including a native or chimeric immunoglobulin variable region, a toxin such as pseudomonas exotoxin or ricin, or a polypeptide therapeutic agent not otherwise normally associated with a constant domain; and V_L , V_H , C_L and C_H represent light or heavy chain variable or constant domains of an immunoglobulin. These diagrams are understood to be merely exemplary of general assembled immunoglobulin structures, and do not encompass all possibilities. It will be understood, for example, that there might desirably be several different "A"s or "X"s in any of these constructs.

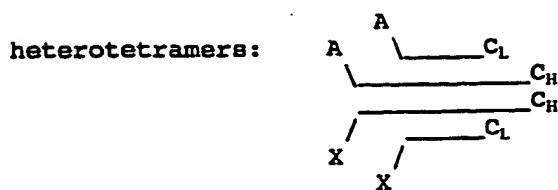
monomer: A _____ C_L or C_H

homodimer: A _____ C_L or C_H
 A _____ C_L or C_H

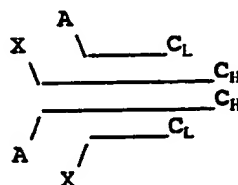
heterodimer: A _____ C_L or C_H
 X _____ C_L or C_H

homotetramer: A _____ C_L
 A _____ C_H
 A _____ C_H
 A _____ C_L

-12-

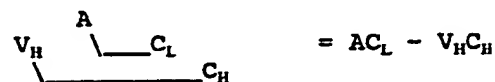
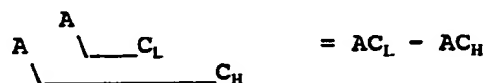
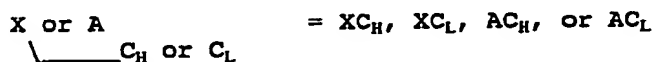


and



It will be understood that these diagrams are merely illustrative, and that the chains of the multimers are believed to be disulfide bonded in the same fashion as native immunoglobulins. According to this invention, hybrid immunoglobulins are readily secreted from mammalian cells transformed with the appropriate nucleic acid. The secreted forms include those wherein the binding partner epitope is present in heavy chain dimers, light chain monomers or dimers, and heavy and light chain heterotetramers wherein the binding partner epitope is present fused to one or more light or heavy chains, including heterotetramers wherein up to and including all four variable region analogues are substituted. Where a light-heavy chain non-binding partner variable-like domain is present, a heterofunctional antibody thus is provided.

Chains or basic units of varying structure may be utilized to assemble the monomers and hetero- and homo-multimers and immunoglobulins of this invention. Specific examples of these basic units are diagrammed below and their equivalents (for purposes of the attenuated formulae *infra*) are indicated.



-13-



5 Various exemplary assembled novel immunoglobulins produced in accordance with this invention are schematically diagrammed below. In addition to the symbols defined above, n is an integer, and Y designates a covalent cross-linking moiety.

- (a) AC_L ;
- (b) $\text{AC}_L\text{-AC}_L$;
- 10 (c) $\text{AC}_H\text{-[AC}_H, \text{AC}_L\text{-AC}_H, \text{AC}_L\text{-V}_H\text{C}_H, \text{V}_L\text{C}_L\text{-AC}_H, \text{V}_L\text{C}_L\text{-V}_H\text{C}_H, \text{XC}_H, \text{XC}_L, \text{XC}_L\text{-XC}_H, \text{XC}_L\text{-V}_H\text{C}_H, \text{XC}_H\text{-V}_L\text{C}_L, \text{XC}_L\text{-AC}_H, \text{or } \text{AC}_L\text{-XC}_H]$;
- (d) $\text{AC}_L\text{-AC}_H\text{-[AC}_H, \text{AC}_L\text{-AC}_H, \text{AC}_L\text{-V}_H\text{C}_H, \text{V}_L\text{C}_L\text{-AC}_H, \text{V}_L\text{C}_L\text{-V}_H\text{C}_H, \text{XC}_H, \text{XC}_L, \text{XC}_L\text{-XC}_H, \text{XC}_L\text{-V}_H\text{C}_H, \text{XC}_H\text{-V}_L\text{C}_L, \text{XC}_L\text{-AC}_H, \text{or } \text{AC}_L\text{-XC}_H]$;
- (e) $\text{AC}_L\text{-V}_H\text{C}_H\text{-[AC}_H, \text{AC}_L\text{-AC}_H, \text{AC}_L\text{-V}_H\text{C}_H, \text{V}_L\text{C}_L\text{-AC}_H, \text{V}_L\text{C}_L\text{-V}_H\text{C}_H, \text{XC}_H, \text{XC}_L, \text{XC}_L\text{-XC}_H, \text{XC}_L\text{-V}_H\text{C}_H, \text{XC}_H\text{-V}_L\text{C}_L, \text{XC}_L\text{-AC}_H, \text{or } \text{AC}_L\text{-XC}_H]$;
- 15 (f) $\text{V}_L\text{C}_L\text{-AC}_H\text{-[AC}_H, \text{AC}_L\text{-AC}_H, \text{AC}_L\text{-V}_H\text{C}_H, \text{V}_L\text{C}_L\text{-AC}_H, \text{V}_L\text{C}_L\text{-V}_H\text{C}_H, \text{XC}_H, \text{XC}_L, \text{XC}_L\text{-XC}_H, \text{XC}_L\text{-V}_H\text{C}_H, \text{XC}_H\text{-V}_L\text{C}_L, \text{XC}_L\text{-AC}_H, \text{or } \text{AC}_L\text{-XC}_H]$;
- (g) $[\text{A-Y}]_n\text{-[V}_L\text{C}_L\text{-V}_H\text{C}_H]_2$;
- (h) XC_H or $\text{XC}_L\text{-[AC}_H, \text{AC}_L\text{-AC}_H, \text{AC}_L\text{-V}_H\text{C}_H, \text{V}_L\text{C}_L\text{-AC}_H, \text{XC}_L\text{-AC}_H, \text{or } \text{AC}_L\text{-XC}_H]$;
- 20 (i) $\text{XC}_L\text{-XC}_H\text{-[AC}_H, \text{AC}_L\text{-AC}_H, \text{AC}_L\text{-V}_H\text{C}_H, \text{V}_L\text{C}_L\text{-AC}_H, \text{XC}_L\text{-AC}_H, \text{or } \text{AC}_L\text{-XC}_H]$;
- (j) $\text{XC}_L\text{-V}_H\text{C}_H\text{-[AC}_H, \text{AC}_L\text{-AC}_H, \text{AC}_L\text{-V}_H\text{C}_H, \text{V}_L\text{C}_L\text{-AC}_H, \text{XC}_L\text{-AC}_H, \text{or } \text{AC}_L\text{-XC}_H]$;
- (k) $\text{XC}_H\text{-V}_L\text{C}_L\text{-[AC}_H, \text{AC}_L\text{-AC}_H, \text{AC}_L\text{-V}_H\text{C}_H, \text{V}_L\text{C}_L\text{-AC}_H, \text{XC}_L\text{-AC}_H, \text{or } \text{AC}_L\text{-XC}_H]$;
- (l) $\text{XC}_L\text{-AC}_H\text{-[AC}_H, \text{AC}_L\text{-AC}_H, \text{AC}_L\text{-V}_H\text{C}_H, \text{V}_L\text{C}_L\text{-AC}_H, \text{V}_L\text{C}_L\text{-V}_H\text{C}_H, \text{XC}_H, \text{XC}_L, \text{XC}_L\text{-XC}_H, \text{XC}_L\text{-V}_H\text{C}_H, \text{XC}_H\text{-V}_L\text{C}_L, \text{XC}_L\text{-AC}_H, \text{or } \text{AC}_L\text{-XC}_H]$;
- 25 (m) $\text{AC}_L\text{-XC}_H\text{-[AC}_H, \text{AC}_L\text{-AC}_H, \text{AC}_L\text{-V}_H\text{C}_H, \text{V}_L\text{C}_L\text{-AC}_H, \text{V}_L\text{C}_L\text{-V}_H\text{C}_H, \text{XC}_H, \text{XC}_L, \text{XC}_L\text{-XC}_H, \text{XC}_L\text{-V}_H\text{C}_H, \text{XC}_H\text{-V}_L\text{C}_L, \text{XC}_L\text{-AC}_H, \text{or } \text{AC}_L\text{-XC}_H]$;

A, X, V or C may be modified with a covalent cross-linking moiety (Y) so to be $(\text{A} - \text{Y})_n$, $(\text{X} - \text{Y})_n$ etc.

30 The ligand binding partner A may also be a multi-chain molecule, e.g. having chains arbitrarily denoted as A_α and A_β . These chains as a unit are located at the sites noted for the single chain "A" above. One of the multiple chains is fused to one immunoglobulin chain (with the remaining chains covalently or noncovalently associated with the fused chain in the normal fashion) or, when the ligand binding partner contains two chains, one chain is separately fused to an immunoglobulin light chain and the other chain to an immunoglobulin heavy chain.

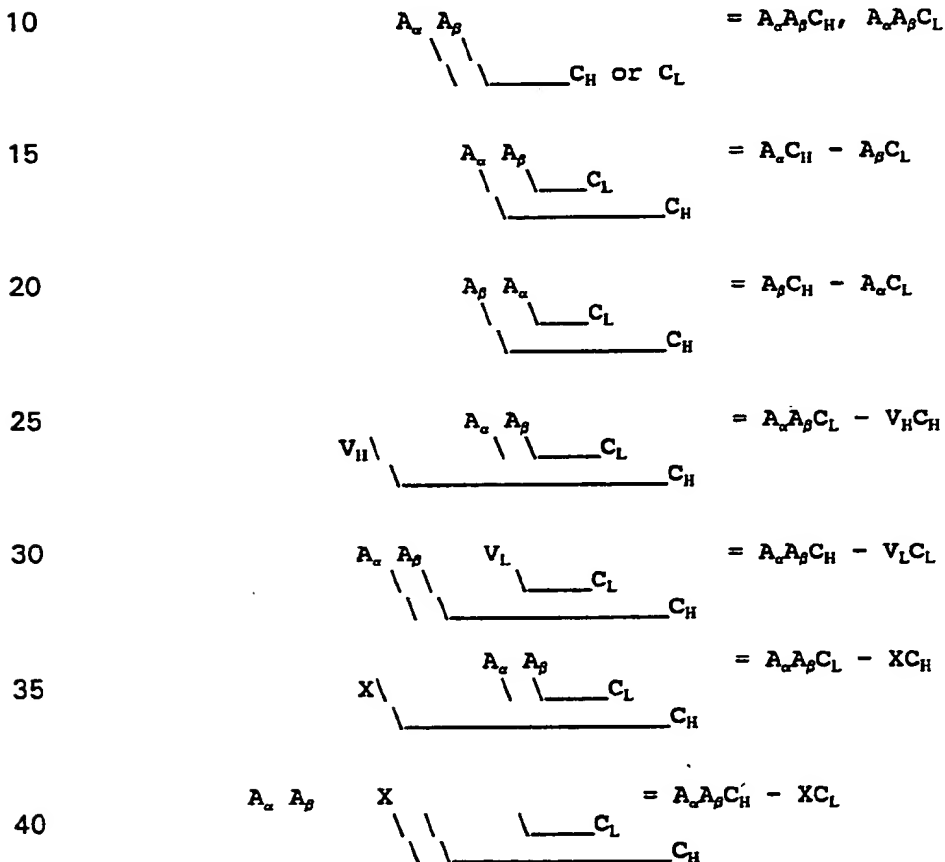
35 It is presently preferred that only one chain of the ligand binding partner be fused to the stable plasma protein. In this case, a fusion through a peptide bond is made between one of the binding partner chains and the stable plasma protein, while the other chain(s) of the ligand binding partner are allowed to associate with the fused chain in the fashion in which

-14-

they associate in nature, *e.g.* by disulfide bonding or hydrophobic interaction. The ligand binding partner chain chosen for peptidyl fusion should be the chain which contains a transmembrane domain, and the fusion will be located substantially adjacent N-terminally from the transmembrane domain or in place of the transmembrane and cytoplasmic domains.

5 Ordinarily, if multiple transmembrane domains are present then one is selected for fusion (or deletion and then fusion) while the other remains unfused or is deleted.

Basic units having the structures as diagrammed below are examples of those used to create monomers, and hetero- and homo-multimers, particularly dimers and trimers with multi-chain ligand binding partners:



Various exemplary novel assembled antibodies having a two-chain ligand binding partner ("A_α and A_β") utilized in unit structures as above are schematically diagrammed below.

- 45 (n) $A_\alpha A_\beta C_L$ -[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, V_LC_L-V_HC_H, XC_H, XC_L, XC_L-XC_H, XC_L-V_HC_H, XC_H-V_LC_L, XC_L-AC_H, AC_L-XC_H, A_αA_βC_H, A_αA_βC_L, A_αC_L-A_βC_H, A_βC_L-A_αC_H, A_αA_βC_L-V_HC_H, A_αA_βC_H-V_LC_L, A_αA_βC_L-XC_H, or A_αA_βC_H-XC_L];
- (o) $A_\alpha A_\beta C_H$ -[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, V_LC_L-V_HC_H, XC_H, XC_L, XC_L-XC_H, XC_L-V_HC_H, XC_H-V_LC_L, XC_L-AC_H, AC_L-XC_H, A_αA_βC_H, A_αA_βC_L, A_αC_L-A_βC_H, A_βC_L-A_αC_H, A_αA_βC_L-V_HC_H, A_αA_βC_H-V_LC_L, A_αA_βC_L-XC_H, or A_αA_βC_H-XC_L];
- 50

- (p) $A_{\alpha}C_L-A_{\beta}C_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, V_LC_L-V_HC_H, XC_H, XC_L, XC_L-XC_H, XC_L-V_HC_H, XC_H-V_LC_L, XC_L-AC_H, AC_L-XC_H, A_{\alpha}A_{\beta}C_H, A_{\alpha}A_{\beta}C_L, A_{\alpha}C_L-A_{\beta}C_H, A_{\beta}C_L-A_{\alpha}C_H, A_{\alpha}A_{\beta}C_L-V_HC_H, A_{\alpha}A_{\beta}C_H-V_LC_L, A_{\alpha}A_{\beta}C_L-XC_H, \text{ or } A_{\alpha}A_{\beta}C_H-XC_L]$;
- (q) $A_{\beta}C_L-A_{\alpha}C_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, V_LC_L-V_HC_H, XC_H, XC_L, XC_L-XC_H, XC_L-V_HC_H, XC_H-V_LC_L, XC_L-AC_H, AC_L-XC_H, A_{\alpha}A_{\beta}C_H, A_{\alpha}A_{\beta}C_L, A_{\alpha}C_L-A_{\beta}C_H, A_{\beta}C_L-A_{\alpha}C_H, A_{\alpha}A_{\beta}C_L-V_HC_H, A_{\alpha}A_{\beta}C_H-V_LC_L, A_{\alpha}A_{\beta}C_L-XC_H, \text{ or } A_{\alpha}A_{\beta}C_H-XC_L]$;
- (r) $A_{\alpha}A_{\beta}C_L-V_HC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, V_LC_L-V_HC_H, XC_H, XC_L, XC_L-XC_H, XC_L-V_HC_H, XC_H-V_LC_L, XC_L-AC_H, AC_L-XC_H, A_{\alpha}A_{\beta}C_H, A_{\alpha}A_{\beta}C_L, A_{\alpha}C_L-A_{\beta}C_H, A_{\beta}C_L-A_{\alpha}C_H, A_{\alpha}A_{\beta}C_L-V_HC_H, A_{\alpha}A_{\beta}C_H-V_LC_L, A_{\alpha}A_{\beta}C_L-XC_H, \text{ or } A_{\alpha}A_{\beta}C_H-XC_L]$;
- (s) $A_{\alpha}A_{\beta}C_H-V_LC_L-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, V_LC_L-V_HC_H, XC_H, XC_L, XC_L-XC_H, XC_L-V_HC_H, XC_H-V_LC_L, XC_L-AC_H, AC_L-XC_H, A_{\alpha}A_{\beta}C_H, A_{\alpha}A_{\beta}C_L, A_{\alpha}C_L-A_{\beta}C_H, A_{\beta}C_L-A_{\alpha}C_H, A_{\alpha}A_{\beta}C_L-V_HC_H, A_{\alpha}A_{\beta}C_H-V_LC_L, A_{\alpha}A_{\beta}C_L-XC_H, \text{ or } A_{\alpha}A_{\beta}C_H-XC_L]$;
- (t) $A_{\alpha}A_{\beta}C_L-XC_H-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, V_LC_L-V_HC_H, XC_H, XC_L, XC_L-XC_H, XC_L-V_HC_H, XC_H-V_LC_L, XC_L-AC_H, AC_L-XC_H, A_{\alpha}A_{\beta}C_H, A_{\alpha}A_{\beta}C_L, A_{\alpha}C_L-A_{\beta}C_H, A_{\beta}C_L-A_{\alpha}C_H, A_{\alpha}A_{\beta}C_L-V_HC_H, A_{\alpha}A_{\beta}C_H-V_LC_L, A_{\alpha}A_{\beta}C_L-XC_H, \text{ or } A_{\alpha}A_{\beta}C_H-XC_L]$;
- (u) $A_{\alpha}A_{\beta}C_H-XC_L-[AC_H, AC_L-AC_H, AC_L-V_HC_H, V_LC_L-AC_H, V_LC_L-V_HC_H, XC_H, XC_L, XC_L-XC_H, XC_L-V_HC_H, XC_H-V_LC_L, XC_L-AC_H, AC_L-XC_H, A_{\alpha}A_{\beta}C_H, A_{\alpha}A_{\beta}C_L, A_{\alpha}C_L-A_{\beta}C_H, A_{\beta}C_L-A_{\alpha}C_H, A_{\alpha}A_{\beta}C_L-V_HC_H, A_{\alpha}A_{\beta}C_H-V_LC_L, A_{\alpha}A_{\beta}C_L-XC_H, \text{ or } A_{\alpha}A_{\beta}C_H-XC_L]$;

The structures shown in the above tables show only key features, e. g. they do not show joining (J) or other domains of the immunoglobulins, nor are disulfide bonds shown. These are omitted in the interests of brevity. However, where such domains are required for binding activity they shall be construed as being present in the ordinary locations which they occupy in the binding partner or immunoglobulin molecules as the case may be.

Where an immunoglobulin V_LV_H antibody combining site is designated above, or where XC_L or XC_H is indicated and X is an immunoglobulin variable region, it preferably is capable of binding to a predetermined antigen. Suitable immunoglobulin combining sites and fusion partners are obtained from IgG-1, -2, -3, or -4 subtypes, IgA, IgE, IgD or IgM, but preferably IgG-1. The schematic examples above are representative of divalent antibodies; more complex structures would result by employing immunoglobulin heavy chain sequences from other classes, e.g. IgM.

A particularly preferred embodiment is a fusion of an N-terminal portion of a LHR, which contains the binding site for the endothelium of lymphoid tissue, to the C-terminal Fc portion of an antibody, containing the effector functions of immunoglobulin G_1 . There are two preferred embodiments of this sort; in one, the entire heavy chain constant region is fused to a portion of the LHR; in another, a sequence beginning in the hinge region just upstream of the papain cleavage site which defines IgG Fc chemically (residue 216, taking the first residue of heavy chain constant region to be 114 [Kabat *et al.*, "Sequences of Proteins of Immunological Interest" 4th Ed., 1987], or analogous sites of other

immunoglobulins) is fused to a portion of the LHR. The latter embodiment is described in the Example 4.

Those compositions of this invention, particularly those in which a biologically active portion of a ligand binding partner is substituted for the variable region of an immunoglobulin chain, are believed to exhibit improved *in vivo* plasma half life. These hybrid immunoglobulins are constructed in a fashion similar to the constructs where a immunoglobulin-like domain is substituted for the variable domain of an immunoglobulin, see e.g. Capon *et al.*, Nature 337:525-531, 1989, Traunecker *et al.*, Nature 339:68-70, 1989, Gascoigne *et al.*, Proc. Nat. Acad. Sci. 84:2936-2940, 1987, and published European application EPO 0 325 224 A2.

The hybrid immunoglobulins of this invention are also constructed in a fashion similar to chimeric antibodies in which a variable domain from an antibody of one species is substituted for the variable domain of another species. See, for example, EP 0 125 023; Munro, Nature 312: (13 December 1984); Neuberger *et al.*, Nature 312: (13 December 1984); Sharon *et al.*, Nature 309: (24 May 1984); Morrison *et al.*, Proc. Natl. Acad. Sci. USA 81:6851-6855 (1984); Morrison *et al.* Science 229:1202-1207 (1985); and Boulian *et al.*, Nature 312:643-646 (13 December 1984). The DNA encoding the binding partner is cleaved by a restriction enzyme at or proximal to the 3' end of the DNA encoding the desired binding partner domain(s) and at a point at or near the DNA encoding the N-terminal end of the mature polypeptide (where use of a different leader is contemplated) or at or proximal to the N-terminal coding region for the binding partner (where a native signal is employed). This DNA fragment then is readily inserted into DNA encoding an immunoglobulin light or heavy chain constant region and, if necessary, tailored by deletional mutagenesis. Preferably, this is a human immunoglobulin when the variant is intended for *in vivo* therapy for humans.

The LHR extracellular domain generally is fused at its C-terminus to the immunoglobulin constant region. The precise site at which the fusion is made is not critical; other sites neighboring or within the extracellular region may be selected in order to optimize the secretion or binding characteristics of the soluble LHR-Ig fusion. The optimal site will be determined by routine experimentation. The fusion may typically take the place of either or both the transmembrane and cytoplasmic domains.

DNA encoding immunoglobulin light or heavy chain constant regions is known or readily available from cDNA libraries or is synthesized. See for example, Adams *et al.*, Biochemistry 19:2711-2719 (1980); Gough *et al.*, Biochemistry 19:2702-2710 (1980); Dolby *et al.*, P.N.A.S. USA, 77:6027-6031 (1980); Rice *et al.*, P.N.A.S. USA 79:7862-7865 (1982); Falkner *et al.*, Nature 298:286-288 (1982); and Morrison *et al.*, Ann. Rev. Immunol. 2:239-256 (1984).

DNA sequences encoding the LHR are provided herein. DNA sequences encoding other desired binding partners which are known or readily available from cDNA libraries are suitable in the practice of this invention.

DNA encoding a fusion of this invention is transfected into a host cell for expression. If multimers are desired then the host cell is transformed with DNA encoding each chain that will make up the multimer, with the host cell optimally being selected to be capable of assembling the chains of the multimers in the desired fashion. If the host cell is producing an immunoglobulin prior to transfection then one need only transfect with the binding partner fused to light or to heavy chain to produce a heteroantibody. The aforementioned immunoglobulins having one or more arms bearing the binding partner domain and one or more arms bearing companion variable regions result in dual specificity for the binding partner ligand and for an antigen or therapeutic moiety. Multiply cotransformed cells are used with the above-described recombinant methods to produce polypeptides having multiple specificities, such as the heterotetrameric immunoglobulins discussed above.

In general, it has been found that the fusions are expressed intracellularly and well secreted, but a great deal of variation is routinely encountered in the degree of secretion of various fusions from recombinant hosts.

Additionally, procedures are known for producing intact heteroantibodies from immunoglobulins having different specificities. These procedures are adopted for the *in vitro* synthesis or production of heterochimeric antibodies by simply substituting the binding partner-immunoglobulin chains where an immunoglobulin or immunoglobulin hybrid was previously used.

In an alternative method for producing a heterofunctional antibody, host cells producing a binding partner-immunoglobulin fusion, e.g. transfected myelomas, also are fused with B cells or hybridomas which secrete antibody having the desired companion specificity for an antigen. Heterobifunctional antibody is recovered from the culture medium of such hybridomas, and thus may be produced somewhat more conveniently than by conventional *in vitro* resorting methods (EP 68,763).

This invention also contemplates amino acid sequence variants of the LHR or other binding partner. Amino acid sequence variants of the binding partner are prepared with various objectives in mind, including increasing the affinity of the binding partner for its ligand, facilitating the stability, purification and preparation of the binding partner, modifying its plasma half life, improving therapeutic efficacy, and lessening the severity or occurrence of side effects during therapeutic use of the binding partner. In the discussion below, amino acid sequence variants of the LHR are provided, exemplary of the variants that may be selected for other ligand binding partners.

Amino acid sequence variants of the ligand binding partner fall into one or more of three classes: Insertional, substitutional, or deletional variants. These variants ordinarily are prepared by site-specific mutagenesis of nucleotides in the DNA encoding the ligand binding partner, by which DNA encoding the variant is obtained, and thereafter expressing the DNA in recombinant cell culture. However, fragments having up to about 100-150 amino acid

residues are prepared conveniently by *in vitro* synthesis. While the following discussion in part refers to LHR, it applies with equal effect to any ligand binding partner to the extent it is applicable to the structure or function thereof.

5 The amino acid sequence variants of the LHR are predetermined variants not found in nature or naturally occurring alleles. The LHR variants typically exhibit the same qualitative biological—for example, ligand binding—activity as the naturally occurring HuLHR or MLHR analogue. However, the LHR variants and derivatives that are not capable of binding to their ligands are useful nonetheless (a) as a reagent in diagnostic assays for the LHR or antibodies to the LHR, (b) when insolubilized in accord with known methods, as agents for purifying anti-
10 LHR antibodies from antisera or hybridoma culture supernatants, and (c) as immunogens for raising antibodies to the LHR or as immunoassay kit components (labelled, as a competitive reagent for the native LHR or unlabelled as a standard for the LHR assay) so long as at least one LHR epitope remains active.

While the site for introducing an amino acid sequence variation is predetermined, the
15 mutation *per se* need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random or saturation mutagenesis (where all 20 possible residues are inserted) is conducted at the target codon and the expressed LHR variant is screened for the optimal combination of desired activities. Such screening is within the ordinary skill in the art.

20 Amino acid insertions usually will be on the order of about from 1 to 10 amino acid residues; substitutions are typically introduced for single residues; and deletions will range about from 1 to 30 residues. Deletions or insertions preferably are made in adjacent pairs, i.e. a deletion of 2 residues or insertion of 2 residues. It will be amply apparent from the following discussion that substitutions, deletions, insertions or any combination thereof are
25 introduced or combined to arrive at a final construct.

Insertional amino acid sequence variants of the LHR are those in which one or more amino acid residues extraneous to the LHR are introduced into a predetermined site in the target LHR and which displace the preexisting residues.

Commonly, insertional variants are fusions of heterologous proteins or polypeptides to
30 the amino or carboxyl terminus of the LHR. Such variants are referred to as fusions of the LHR and a polypeptide containing a sequence which is other than that which is normally found in the LHR at the inserted position. Several groups of fusions are contemplated herein.

The novel polypeptides of this invention are useful in diagnostics or in purification of the ligand binding partner by immunoaffinity techniques known *per se*. Alternatively, in the
35 purification of the binding partner, the novel polypeptides are used to adsorb the fusion from impure admixtures, after which the fusion is eluted and, if desired, the binding partner is recovered from the fusion, e.g. by enzymatic cleavage.

Desirable fusions of the binding partner, which may or may not also be immunologically active, include fusions of the mature binding partner sequence with a signal sequence heterologous to the binding partner.

In the case of the LHR, and where desired with other selected binding proteins, signal sequence fusions are employed in order to more expeditiously direct the secretion of the LHR. The heterologous signal replaces the native LHR signal, and when the resulting fusion is recognized, i.e. processed and cleaved by the host cell, the LHR is secreted. Signals are selected based on the intended host cell, and may include bacterial yeast, mammalian and viral sequences. The native LHR signal or the herpes gD glycoprotein signal is suitable for use in mammalian expression systems.

Substitutional variants are those in which at least one residue in the Fig. 1 or 2 sequence has been removed and a different residue inserted in its place. Such substitutions generally are made in accordance with the following Table 1 when it is desired to finely modulate the characteristics of the LHR.

TABLE 1

	<u>Original Residue</u>	<u>Exemplary Substitutions</u>
	Ala	ser
	Arg	lys
	Asn	gln; his
20	Asp	glu
	Cys	ser; ala
	Gln	asn
	Glu	asp
	Gly	pro
25	His	asn; gln
	Ile	leu; val
	Leu	ile; val
	Lys	arg; gln; glu
	Met	leu; ile
30	Phe	met; leu; tyr
	Ser	thr
	Thr	ser
	Trp	tyr
	Tyr	trp; phe
35	Val	ile; leu

Novel amino acid sequences, as well as isosteric analogs (amino acid or otherwise), as included within the scope of this invention.

Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative than those in Table 1, i.e., selecting residues that differ more significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site or (c) the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in LHR properties will be those in which (a) a hydrophilic residue, e.g. seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having a side chain, e.g., glycine.

Some deletions, insertions, and substitutions will not produce radical changes in the characteristics of the LHR molecule. However, when it is difficult to predict the exact effect of the substitution, deletion, or insertion in advance of doing so, for example when modifying the LHR carbohydrate binding domain or an immune epitope, one skilled in the art will appreciate that the effect will be evaluated by routine screening assays. For example, a variant typically is made by site specific mutagenesis of the LHR-encoding nucleic acid, expression of the variant nucleic acid in recombinant cell culture and, optionally, purification from the cell culture for example by immunoaffinity adsorption on a polyclonal anti-LHR column (in order to adsorb the variant by at least one remaining immune epitope). The activity of the cell lysate or purified LHR variant is then screened in a suitable screening assay for the desired characteristic. For example, a change in the immunological character of the LHR, such as affinity for a given antibody such as Mel-14, is measured by a competitive-type immunoassay. As more becomes known about the functions *in vivo* of the LHR other assays will become useful in such screening. Modifications of such protein properties as redox or thermal stability, hydrophobicity, susceptibility to proteolytic degradation, or the tendency to aggregate with carriers or into multimers are assayed by methods well known to the artisan.

Substitutional variants of the LHR also include variants where functionally homologous domains of other proteins are substituted by routine methods for one or more of the above-identified LHR domains. Where the variant is a fragment of a particular domain of the LHR, it preferably but not necessarily has at least ~70% homology to the corresponding LHR domain as defined herein. Figs. 6A-6C may be used by those skilled in the art for sources for such substitutable domains. For example, the flesh fly lectin whose sequence is shown in Fig. 6A may be modified to rise to the level of at least ~70% homology with the carbohydrate binding domain of the LHR, and then substituted for that domain. Similarly, coagulation Factor X, whose sequence is shown in Fig. 6B may be modified to rise to the level of at least ~70% homology with the egf-domain of the LHR, and then substituted for

that domain. Similar substitutions may desirably be made for the signal sequence, the complement binding domain, the transmembrane domain, and for the cytoplasmic domain.

Another class of LHR variants are deletional variants. Deletions are characterized by the removal of one or more amino acid residues from the LHR sequence. Typically, the transmembrane and cytoplasmic domains, or only the cytoplasmic domains of the LHR are deleted. However, deletion from the LHR C-terminal to any other suitable site N-terminal to the transmembrane region which preserves the biological activity or immune cross-reactivity of the LHR is suitable. Excluded from the scope of deletional variants are the protein digestion fragments heretofore obtained in the course of elucidating amino acid sequences of the LHR, and protein fragments having less than ~70% sequence homology to any of the above-identified LHR domains.

Immunoglobulin fusions may be made with fragments of the LHR, such as the complement binding domain, the carbohydrate domain, and the epidermal growth factor domain. The complement binding domain fusion finds usefulness in the diagnosis and treatment of complement-mediated diseases, as well as in the oligomerization of the fusion with the LHR or with other components on the lymphocyte surface.

Deletions of cysteine or other labile residues also may be desirable, for example in increasing the oxidative stability of the LHR. Deletion or substitutions of potential proteolysis sites, e.g. Arg Arg, is accomplished by deleting one of the basic residues or substituting one by glutaminy or histidyl residues.

In one embodiment, the LHR is comprised of the carbohydrate binding domain in the absence of a complement binding domain and/or the egf domain. This embodiment may or may not contain either or both the transmembrane and cytoplasmic regions.

A preferred class of substitutional or deletional variants are those involving a transmembrane region of the LHR. Transmembrane regions of LHR subunits are highly hydrophobic or lipophilic domains that are the proper size to span the lipid bilayer of the cellular membrane. They are believed to anchor the LHR in the cell membrane, and allow for homo- or heteropolymeric complex formation with the LHR.

Inactivation of the transmembrane domain of the LHR and any other binding partner where one is present, typically by deletion or substitution of transmembrane domain hydroxylation residues, will facilitate recovery and formulation by reducing its cellular or membrane lipid affinity and improving its aqueous solubility. If the transmembrane and cytoplasmic domains are deleted one avoids the introduction of potentially immunogenic epitopes, either by exposure of otherwise intracellular polypeptides that might be recognized by the body as foreign or by insertion of heterologous polypeptides that are potentially immunogenic. Inactivation of the membrane binding function is accomplished by deletion of sufficient residues to produce a substantially hydrophilic hydropathy profile at this site or by substituting with heterologous residues which accomplish the same result.

A principal advantage of the transmembrane inactivated LHR is that it may be secreted into the culture medium of recombinant hosts. This variant is soluble in body fluids such as blood and does not have an appreciable affinity for cell membrane lipids, thus considerably simplifying its recovery from recombinant cell culture.

5 As a general proposition, all variants will not have a functional transmembrane domain and preferably will not have a functional cytoplasmic sequence.

For example, the transmembrane domain may be substituted by any amino acid sequence, e.g. a random or predetermined sequence of about 5 to 50 serine, threonine, lysine, arginine, glutamine, aspartic acid and like hydrophilic residues, which altogether exhibit
10 a hydrophilic hydropathy profile. Like the deletional (truncated) LHR, these variants are secreted into the culture medium of recombinant hosts.

Examples of HuLHR amino acid sequence variants are described in the table below. The residue following the residue number indicates the replacement or inserted amino acids.

Table 2

15	<u>Substitutions</u>	<u>Deletions</u>
	Arg58-Asp59: Lys-Glu	Gly96-Ile97
	Ala71: Ser	Asn136
	Lys78: Gln	Ser166
	Asp116: Glu	Ser220
20	Leu150: Val	Asn271
	His168: Gln	Ile296
	Ile174: Leu	
	Asn181: Gln	
	Thr211: Ser	<u>Insertions</u>
25	Phe214: Leu	67-Glu-Ser-Ala
	Ser226: Thr	83-Gly-Thr-Thr
	Phe244: Met	209-Asn
	Thr282: Ser	241-Val-Glu-Asn
	Ile288: Val	292-Tyr-Tyr-Tyr
30	Lys298-Lys299: Arg-Arg	
	Ile302: Leu	

35 Preferably, the variants represent conservative substitutions. It will be understood that some variants may exhibit reduced or absent biological activity. These variants nonetheless are useful as standards in immunoassays for the LHR so long as they retain at least one immune epitope of the LHR.

40 Glycosylation variants are included within the scope of the HuLHR. They include variants completely lacking in glycosylation (unglycosylated) and variants having at least one less glycosylated site than the native form (deglycosylated) as well as variants in which the glycosylation has been changed. Included are deglycosylated and unglycosylated amino acid sequence variants, deglycosylated and unglycosylated LHR having the native, unmodified amino acid sequence of the LHR, and other glycosylation variants. For example, substitutional or deletional mutagenesis is employed to eliminate the N- or O-linked

glycosylation sites of the LHR, e.g., the asparagine residue is deleted or substituted for by another basic residue such as lysine or histidine. Alternatively, flanking residues making up the glycosylation site are substituted or deleted, even though the asparagine residues remain unchanged, in order to prevent glycosylation by eliminating the glycosylation recognition site.

5 Additionally, unglycosylated LHR which has the amino acid sequence of the native LHR is produced in recombinant prokaryotic cell culture because prokaryotes are incapable of introducing glycosylation into polypeptides.

10 Glycosylation variants are produced by selecting appropriate host cells or by *in vitro* methods. Yeast, for example, introduce glycosylation which varies significantly from that of mammalian systems. Similarly, mammalian cells having a different species (e.g. hamster, murine, insect, porcine, bovine or ovine) or tissue origin (e.g. lung, liver, lymphoid, mesenchymal or epidermal) than the source of the LHR are routinely screened for the ability to introduce variant glycosylation as characterized for example by elevated levels of mannose or variant ratios of mannose, fucose, sialic acid, and other sugars typically found in

15 mammalian glycoproteins. *in vitro* processing of the LHR typically is accomplished by enzymatic hydrolysis, e.g. neuraminidase digestion.

20 Covalent modifications of the LHR molecule are included within the scope hereof. Such modifications are introduced by reacting targeted amino acid residues of the recovered protein with an organic derivatizing agent that is capable of reacting with selected side chains or terminal residues, or by harnessing mechanisms of post-translational modification that function in selected recombinant host cells. The resulting covalent derivatives are useful in programs directed at identifying residues important for biological activity, for immunoassays of the LHR or for the preparation of anti-LHR antibodies for immunoaffinity purification of the recombinant LHR. For example, complete inactivation of the biological activity of the protein

25 after reaction with ninhydrin would suggest that at least one arginyl or lysyl residue is critical for its activity, whereafter the individual residues which were modified under the conditions selected are identified by isolation of a peptide fragment containing the modified amino acid residue. Such modifications are within the ordinary skill in the art and are performed without undue experimentation.

30 Derivatization with bifunctional agents is useful for preparing intermolecular aggregates of the hybrid immunoglobulin with polypeptides as well as for cross-linking the hybrid immunoglobulin to a water insoluble support matrix or surface for use in the assay or affinity purification of its ligands. In addition, a study of intrachain cross-links will provide direct information on conformational structure. Commonly used cross-linking agents include 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example

35 esters with 4-azidosalicylic acid, homobifunctional imidoesters including disuccinimidyl esters such as 3,3'-dithiobis (succinimidyl-propionate), and bifunctional maleimides such as bis-N-maleimido-1,8-octane. Derivatizing agents such as methyl-3-[(p-azido-phenyl)dithio]

propioimide yield photoactivatable intermediates which are capable of forming cross-links in the presence of light. Alternatively, reactive water insoluble matrices such as cyanogen bromide activated carbohydrates and the systems reactive substrates described in U.S. patents 3,959,080; 3,969,287; 3,691,016; 4,195,128; 4,247,642; 4,229,537; 4,055,635; and 4,330,440 are employed for protein immobilization and cross-linking.

Certain post-translational derivatizations are the result of the action of recombinant host cells on the expressed polypeptide. Glutaminyl and asparaginyl residues are frequently post-translationally deamidated to the corresponding glutamyl and aspartyl residues. Alternatively, these residues are deamidated under mildly acidic conditions. Either form of these residues falls within the scope of this invention.

Other post-translational modifications include hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α -amino groups of lysine, arginine, and histidine side chains (T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco pp 79-86 [1983]), acetylation of the N-terminal amine and, in some instances, amidation of the C-terminal carboxyl.

Other derivatives comprise the novel polypeptides of this invention covalently bonded to a nonproteinaceous polymer. The nonproteinaceous polymer ordinarily is a hydrophilic synthetic polymer, i.e., a polymer not otherwise found in nature. However, polymers which exist in nature and are produced by recombinant or *in vitro* methods are useful, as are polymers which are isolated from nature. Hydrophilic polyvinyl polymers fall within the scope of this invention, e.g. polyvinylalcohol and polyvinylpyrrolidone. Particularly useful are polyalkylene ethers such as polyethylene glycol, polypropylene glycol, polyoxyethylene esters or methoxy polyethylene glycol; polyoxyalkylenes such as polyoxyethylene, polyoxypropylene, and block copolymers of polyoxyethylene and polyoxypropylene (Pluronic); polymethacrylates; carbomers; branched or unbranched polysaccharides which comprise the saccharide monomers D-mannose, D- and L-galactose, fucose, fructose, D-xylose, L-arabinose, D-glucuronic acid, sialic acid, D-galacturonic acid, D-mannuronic acid (e.g. polymannuronic acid, or alginic acid), D-glucosamine, D-galactosamine, D-glucose and neuraminic acid including homopolysaccharides and heteropolysaccharides such as lactose, amylopectin, starch, hydroxyethyl starch, amylose, dextran sulfate, dextran, dextrans, glycogen, or the polysaccharide subunit of acid mucopolysaccharides, e.g. hyaluronic acid; polymers of sugar alcohols such as polysorbitol and polymannitol; and heparin or heparan.

Where the polysaccharide is the native glycosylation or the glycosylation attendant on recombinant expression, the site of substitution may be located at other than a native N or O-linked glycosylation site wherein an additional or substitute N or O-linked site has been introduced into the molecule. Mixtures of such polymers may be employed, or the polymer may be homogeneous. The polymer prior to crosslinking need not be, but preferably is, water soluble, but the final conjugate must be water soluble. In addition, the polymer should not

-25-

be highly immunogenic in the conjugate form, nor should it possess viscosity that is incompatible with intravenous infusion or injection if it is intended to be administered by such routes.

5 Preferably the polymer contains only a single group which is reactive. This helps to avoid cross-linking of protein molecules. However, it is within the scope herein to optimize reaction conditions to reduce cross-linking, or to purify the reaction products through gel filtration or chromatographic sieves to recover substantially homogeneous derivatives.

10 The molecular weight of the polymer may desirably range from about 100 to 500,000, and preferably is from about 1,000 to 20,000. The molecular weight chosen will depend upon the nature of the polymer and the degree of substitution. In general, the greater the hydrophilicity of the polymer and the greater the degree of substitution, the lower the molecular weight that can be employed. Optimal molecular weights will be determined by routine experimentation.

15 The polymer generally is covalently linked to the hybrid immunoglobulin herein through a multifunctional crosslinking agent which reacts with the polymer and one or more amino acid or sugar residues of the hybrid. However, it is within the scope of this invention to directly crosslink the polymer by reacting a derivatized polymer with the hybrid, or vice versa.

20 The covalent crosslinking site on the hybrid immunoglobulin includes the N-terminal amino group and epsilon amino groups found on lysine residues, as well as other amino, imino, carboxyl, sulfhydryl, hydroxyl or other hydrophilic groups. The polymer may be covalently bonded directly to the hybrid without the use of a multifunctional (ordinarily bifunctional) crosslinking agent. Covalent bonding to amino groups is accomplished by known chemistries based upon cyanuric chloride, carbonyl diimidazole, aldehyde reactive groups (PEG alkoxide plus diethyl acetal of bromoacetaldehyde; PEG plus DMSO and acetic anhydride, or PEG chloride plus the phenoxide of 4-hydroxybenzaldehyde, succinimidyl active
25 esters, activated dithiocarbonate PEG, 2,4,5-trichlorophenylchloroformate or p-nitrophenylchloroformate activated PEG. Carboxyl groups are derivatized by coupling PEG-amine using carbodiimide.

30 Polymers are conjugated to oligosaccharide groups by oxidation using chemicals, e.g. metaperiodate, or enzymes, e.g. glucose or galactose oxidase, (either of which produces the aldehyde derivative of the carbohydrate), followed by reaction with hydrazide or amino-derivatized polymers, in the same fashion as is described by Heitzmann *et al.*, P.N.A.S., 71:3537-3541 (1974) or Bayer *et al.*, Methods in Enzymology, 62:310 (1979), for the labeling of oligosaccharides with biotin or avidin. Further, other chemical or enzymatic
35 methods which have been used heretofore to link oligosaccharides and polymers are suitable. Substituted oligosaccharides are particularly advantageous because, in general, there are fewer substitutions than amino acid sites for derivatization, and the oligosaccharide products thus will be more homogeneous. The oligosaccharide substituents also are optionally

modified by enzyme digestion to remove sugars, e.g. by neuraminidase digestion, prior to polymer derivatization.

The polymer will bear a group which is directly reactive with an amino acid side chain, or the N- or C- terminus of the polypeptide herein, or which is reactive with the multifunctional cross-linking agent. In general, polymers bearing such reactive groups are known for the preparation of immobilized proteins. In order to use such chemistries here, one should employ a water soluble polymer otherwise derivatized in the same fashion as insoluble polymers heretofore employed for protein immobilization. Cyanogen bromide activation is a particularly useful procedure to employ in crosslinking polysaccharides.

"Water soluble" in reference to the starting polymer means that the polymer or its reactive intermediate used for conjugation is sufficiently water soluble to participate in a derivatization reaction.

"Water soluble" in reference to the polymer conjugate means that the conjugate is soluble in physiological fluids such as blood.

The degree of substitution with such a polymer will vary depending upon the number of reactive sites on the protein, whether all or a fragment of the protein is used, whether the protein is a fusion with a heterologous protein, the molecular weight, hydrophilicity and other characteristics of the polymer, and the particular protein derivatization sites chosen. In general, the conjugate contains about from 1 to 10 polymer molecules, while any heterologous sequence may be substituted with an essentially unlimited number of polymer molecules so long as the desired activity is not significantly adversely affected. The optimal degree of crosslinking is easily determined by an experimental matrix in which the time, temperature and other reaction conditions are varied to change the degree of substitution, after which the ability of the conjugates to function in the desired fashion is determined.

The polymer, e.g. PEG, is crosslinked by a wide variety of methods known *per se* for the covalent modification of proteins with nonproteinaceous polymers such as PEG. Certain of these methods, however, are not preferred for the purposes herein. Cyanuric chloride chemistry leads to many side reactions, including protein cross-linking. In addition, it may be particularly likely to lead to inactivation of proteins containing sulfhydryl groups. Carbonyl diimidazole chemistry (Beauchamp *et al.*, "Anal. Biochem." 131:25-33 [1983]) requires high pH (>8.5), which can inactivate proteins. Moreover, since the "activated PEG" intermediate can react with water, a very large molar excess of "activated PEG" over protein is required. The high concentrations of PEG required for the carbonyl diimidazole chemistry also led to problems with purification, as both gel filtration chromatography and hydrophobic interaction chromatography are adversely effected. In addition, the high concentrations of "activated PEG" may precipitate protein, a problem that *per se* has been noted previously (Davis, U.S. Patent 4,179,337). On the other hand, aldehyde chemistry (Royer, U.S. Patent 4,002,531) is more efficient since it requires only a 40 fold molar excess of PEG and a 1-2 hr incubation.

However, the manganese dioxide suggested by Royer for preparation of the PEG aldehyde is problematic "because of the pronounced tendency of PEG to form complexes with metal-based oxidizing agents" (Harris *et al.*, "J. Polym. Sci., Polym. Chem. Ed." 22:341-352 [1984]). The use of a Moffatt oxidation, utilizing DMSO and acetic anhydride, obviates this problem. In addition, the sodium borohydride suggested by Royer must be used at a high pH and has a significant tendency to reduce disulfide bonds. In contrast, sodium cyanoborohydride, which is effective at neutral pH and has very little tendency to reduce disulfide bonds is preferred.

The conjugates of this invention are separated from unreacted starting materials by gel filtration. Heterologous species of the conjugates are purified from one another in the same fashion.

The polymer also may be water insoluble, as a hydrophilic gel or a shaped article such as surgical tubing in the form of catheters or drainage conduits.

DNA encoding the LHR and other ligand binding partners is synthesized by *in vitro* methods or is obtained readily from lymphocyte cDNA libraries. The means for synthetic creation of the DNA encoding the LHR, either by hand or with an automated apparatus, are generally known to one of ordinary skill in the art, particularly in light of the teachings contained herein. As examples of the current state of the art relating to polynucleotide synthesis, one is directed to Maniatis *et al.* Molecular Cloning--A Laboratory Manual, Cold Spring Harbor Laboratory (1984), and Horvath *et al.* An Automated DNA Synthesizer Employing Deoxynucleoside 3'-Phosphoramidites, Methods in Enzymology 154: 313-326, 1987.

Alternatively, to obtain DNA encoding the LHR from sources other than murine or human, since the entire DNA sequence for the preferred embodiment of the HuLHR (Fig. 1) and of the MLHR (Fig. 2) are given, one needs only to conduct hybridization screening with labelled DNA encoding either HuLHR or MLHR or fragments thereof (usually, greater than about 20, and ordinarily about 50bp) in order to detect clones which contain homologous sequences in the cDNA libraries derived from the lymphocytes of the particular animal, followed by analyzing the clones by restriction enzyme analysis and nucleic acid sequencing to identify full-length clones. If full length clones are not present in the library, then appropriate fragments are recovered from the various clones and ligated at restriction sites common to the fragments to assemble a full-length clone. DNA encoding the LHR from other animal species is obtained by probing libraries from such species with the human or murine sequences, or by synthesizing the genes *in vitro* DNA for other binding partners having known sequence may be obtained with the use of analogous routine hybridization procedures.

Provided herein are nucleic acid sequences that hybridize under stringent conditions to a fragment of the DNA sequence in Fig. 1 or Fig. 2, which fragment is greater than about 10 bp, preferably 20-50 bp, and even greater than 100 bp. Also included within the scope

-28-

hereof are nucleic acid sequences that hybridize under stringent conditions to a fragment of the LHR other than the signal, or transmembrane, or cytoplasmic domains.

Included also within the scope hereof are nucleic acid probes which are capable of hybridizing under stringent conditions to the cDNA of the LHR or to the genomic gene for the LHR (including introns and 5' or 3' flanking regions extending to the adjacent genes or about 5,000 bp, whichever is greater).

Identification of the genomic DNA for the LHR or other binding partner is a straightforward matter of probing a particular genomic library with the cDNA or its fragments which have been labelled with a detectable group, e.g. radiophosphorus, and recovering clone(s) containing the gene. The complete gene is pieced together by "walking" if necessary. Typically, such probes do not encode sequences with less than 70% homology to HuLHR or MLHR, and they range from about 10 to 100 bp in length. Homologies and sizes with respect to other binding partners may be determined without undue experimentation.

In general, prokaryotes are used for cloning of DNA sequences in constructing the vectors useful in the invention. For example, *E. coli* K12 strain 294 (ATCC No. 31446) is particularly useful. Other microbial strains which may be used include *E. coli* B and *E. coli* X1776 (ATCC No. 31537). These examples are illustrative rather than limiting. Alternatively, *in vitro* methods of cloning, e.g. polymerase chain reaction, are suitable.

The polypeptides of this invention are expressed directly in recombinant cell culture as an N-terminal methionyl analogue, or as a fusion with a polypeptide heterologous to the hybrid/portion, preferably a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the hybrid/portion. For example, in constructing a prokaryotic secretory expression vector for the LHR, the native LHR signal is employed with hosts that recognize that signal. When the secretory leader is "recognized" by the host, the host signal peptidase is capable of cleaving a fusion of the leader polypeptide fused at its C-terminus to the desired mature LHR. For host prokaryotes that do not process the LHR signal, the signal is substituted by a prokaryotic signal selected for example from the group of the alkaline phosphatase, penicillinase, lpp or heat stable enterotoxin II leaders. For yeast secretion the human LHR signal may be substituted by the yeast invertase, alpha factor or acid phosphatase leaders. In mammalian cell expression the native signal is satisfactory for mammalian LHR, although other mammalian secretory protein signals are suitable, as are viral secretory leaders, for example the herpes simplex gD signal.

The novel polypeptides may be expressed in any host cell, but preferably are synthesized in mammalian hosts. However, host cells from prokaryotes, fungi, yeast, insects and the like are also used for expression. Exemplary prokaryotes are the strains suitable for cloning as well as *E. coli* W3110 (F⁻ prototrophic, ATCC No. 27325), other enterobacteriaceae such as Serratia marcescans, bacilli and various pseudomonads. Preferably the host cell should secrete minimal amounts of proteolytic enzymes.

-29-

Expression hosts typically are transformed with DNA encoding the hybrid which has been ligated into an expression vector. Such vectors ordinarily carry a replication site (although this is not necessary where chromosomal integration will occur). Expression vectors also include marker sequences which are capable of providing phenotypic selection in transformed cells, as will be discussed further below. For example, *E. coli* is typically transformed using pBR322, a plasmid derived from an *E. coli* species (Bolivar, *et al.* Gene 2: 95 [1977]). pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells, whether for purposes of cloning or expression. Expression vectors also optimally will contain sequences which are useful for the control of transcription and translation, e.g., promoters and Shine-Dalgarno sequences (for prokaryotes) or promoters and enhancers (for mammalian cells). The promoters may be, but need not be, inducible; surprisingly, even powerful constitutive promoters such as the CMV promoter for mammalian hosts have been found to produce the LHR without host cell toxicity. While it is conceivable that expression vectors need not contain any expression control, replicative sequences or selection genes, their absence may hamper the identification of hybrid transformants and the achievement of high level hybrid immunoglobulin expression.

Promoters suitable for use with prokaryotic hosts illustratively include the β -lactamase and lactose promoter systems (Chang *et al.*, "Nature", 275: 615 [1978]; and Goeddel *et al.*, "Nature" 281: 544 [1979]), alkaline phosphatase, the tryptophan (*trp*) promoter system (Goeddel "Nucleic Acids Res." 8: 4057 [1980] and EPO Appln. Publ. No. 36,776) and hybrid promoters such as the *tac* promoter (H. de Boer *et al.*, "Proc. Natl. Acad. Sci. USA" 80: 21-25 [1983]). However, other functional bacterial promoters are suitable. Their nucleotide sequences are generally known, thereby enabling a skilled worker operably to ligate them to DNA encoding the LHR (Siebenlist *et al.*, "Cell" 20: 269 [1980]) using linkers or adaptors to supply any required restriction sites. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding the LHR.

In addition to prokaryotes, eukaryotic microbes such as yeast or filamentous fungi are satisfactory. *Saccharomyces cerevisiae* is the most commonly used eukaryotic microorganism, although a number of other strains are commonly available. The plasmid YRp7 is a satisfactory expression vector in yeast (Stinchcomb, *et al.* Nature 282: 39 [1979]; Kingsman *et al.*, Gene 7: 141 [1979]; Tschemper *et al.* Gene 10: 157 [1980]). This plasmid already contains the *trp1* gene which provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example ATCC no. 44076 or PEP4-1 (Jones, Genetics 85: 12 [1977]). The presence of the *trp1* lesion as a characteristic of the yeast host cell genome then provides an effective environment for detecting transformation by growth in the absence of tryptophan.

Suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase (Hitzeman *et al.*, "J. Biol. Chem." 255: 2073 [1980]) or other

glycolytic enzymes (Hess *et al.*, "J. Adv. Enzyme Reg." 7: 149 [1968]; and Holland, "Biochemistry" 17: 4900 [1978]), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in R. Hitzeman *et al.*, European Patent Publication No. 73,657A.

Expression control sequences are known for eucaryotes. Virtually all eukaryotic genes have an AT-rich region located approximately 25 to 30 bases upstream from the site where transcription is initiated. Another sequence found 70 to 80 bases upstream from the start of transcription of many genes is a CXCAAT region where X may be any nucleotide. At the 3' end of most eukaryotic genes is an AATAAA sequence which may be the signal for addition of the poly A tail to the 3' end of the coding sequence. All of these sequences are inserted into mammalian expression vectors.

Suitable promoters for controlling transcription from vectors in mammalian host cells are readily obtained from various sources, for example, the genomes of viruses such as polyoma virus, SV40, adenovirus, MMV (steroid inducible), retroviruses (e.g. the LTR of HIV), hepatitis-B virus and most preferably cytomegalovirus, or from heterologous mammalian promoters, e.g. the beta actin promoter. The early and late promoters of SV40 are conveniently obtained as an SV40 restriction fragment which also contains the SV40 viral origin of replication. Fiers *et al.*, Nature, 273: 113 (1978). The immediate early promoter of the human cytomegalovirus is conveniently obtained as a HindIII E restriction fragment. Greenaway, P.J. *et al.*, Gene 18: 355-360 (1982).

Transcription of a DNA encoding the hybrid immunoglobulin and/or hybrid portions by higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10-300bp, that act on a promoter to increase its transcription. Enhancers are relatively orientation and position independent having been found 5' (Laimins, L. *et al.*, PNAS 78: 993 [1981]) and 3' (Lusky, M.L., *et al.*, Mol. Cell Bio. 3: 1108 [1983]) to the transcription unit, within an intron (Banerji, J.L. *et al.*, Cell 33: 729 [1983]) as well as within the coding sequence itself (Osborne, T.F., *et al.*, Mol. Cell Bio. 4: 1293 [1984]). Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -fetoprotein and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side

of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

5 Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription which may affect mRNA expression. These regions are transcribed as polyadenylated segments in the untranslated portion of the mRNA encoding the hybrid immunoglobulin. The 3' untranslated regions also include transcription termination sites.

10 Expression vectors may contain a selection gene, also termed a selectable marker. Examples of suitable selectable markers for mammalian cells are dihydrofolate reductase (DHFR), thymidine kinase (TK) or neomycin. When such selectable markers are successfully transferred into a mammalian host cell, the transformed mammalian host cell is able to survive if placed under selective pressure. There are two widely used distinct categories of selective regimes. The first category is based on a cell's metabolism and the use of a mutant cell line
15 which lacks the ability to grow independent of a supplemented media. Two examples are CHO DHFR⁻ cells and mouse LTK⁻ cells. These cells lack the ability to grow without the addition of such nutrients as thymidine or hypoxanthine. Because these cells lack certain genes necessary for a complete nucleotide synthesis pathway, they cannot survive unless the missing nucleotides are provided in a supplemented media. An alternative to supplementing
20 the media is to introduce an intact DHFR or TK gene into cells lacking the respective genes, thus altering their growth requirements. Individual cells which were not transformed with the DHFR or TK gene will not be capable of survival in non supplemented media.

The second category of selective regimes is dominant selection which refers to a selection scheme used in any cell type and does not require the use of a mutant cell line.
25 These schemes typically use a drug to arrest growth of a host cell. Those cells which are successfully transformed with a heterologous gene express a protein conferring drug resistance and thus survive the selection regimen. Examples of such dominant selection use the drugs neomycin (Southern *et al.* J. Molec. Appl. Genet. 1: 327 (1982)), mycophenolic acid (Mulligan *et al.* Science 209: 1422 (1980)) or hygromycin (Sugden *et al.* Mol. Cell. Biol. 5: 410-413 (1985)). The three examples given above employ bacterial genes under
30 eukaryotic control to convey resistance to the appropriate drug G418 or neomycin (geneticin), xgpt (mycophenolic acid) or hygromycin, respectively.

Suitable eukaryotic host cells for expressing the hybrid immunoglobulin include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney
35 line (293 or 293 cells subcloned for growth in suspension culture, Graham, F.L. *et al.*, J. Gen Virol. 36: 59 [1977]); baby hamster kidney cells (BHK, ATCC CCL 10); chinese hamster ovary-cells-DHFR (CHO, Urlaub and Chasin, PNAS (USA) 77: 4216, [1980]); mouse sertoli cells (TM4, Mather, J.P., Biol. Reprod. 23: 243-251 [1980]); monkey kidney cells (CV1 ATCC

CCL 70); african green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor (MMT 060562, ATCC CCL51); and, TRI cells (Mather, J.P. *et al.*, Annals N.Y. Acad. Sci. 383: 44-68 [1982]).

Construction of suitable vectors containing the desired coding and control sequences employ standard ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to form the plasmids required.

For analysis to confirm correct sequences in plasmids constructed, the ligation mixtures are used to transform *E. coli* K12 strain 294 (ATCC 31446) and successful transformants selected by ampicillin or tetracycline resistance where appropriate. Plasmids from the transformants are prepared, analyzed by restriction and/or sequenced by the method of Messing *et al.*, Nucleic Acids Res. 9: 309 (1981) or by the method of Maxam *et al.*, Methods in Enzymology 65: 499 (1980).

Host cells are transformed with the expression vectors of this invention and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants or amplifying the genes encoding the desired sequences. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

The host cells referred to in this disclosure encompass cells in *in vitro* culture as well as cells which are within a host animal.

"Transformation" means introducing DNA into an organism so that the DNA is replicable, either as an extrachromosomal element or by chromosomal integration. Unless indicated otherwise, the method used herein for transformation of the host cells is the method of Graham, F. and van der Eb, A., Virology 52: 456-457 (1973). However, other methods for introducing DNA into cells such as by nuclear injection or by protoplast fusion may also be used. If prokaryotic cells or cells which contain substantial cell wall constructions are used, the preferred method of transfection is calcium treatment using calcium chloride as described by Cohen, F.N. *et al.* Proc. Natl. Acad. Sci. (USA), 69: 2110 (1972).

"Transfection" refers to the introduction of DNA into a host cell whether or not any coding sequences are ultimately expressed. Numerous methods of transfection are known to the ordinarily skilled artisan, for example, CaPO_4 and electroporation. Transformation of the host cell is the indicia of successful transfection.

The novel polypeptide is recovered and purified from recombinant cell cultures by known methods, including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, immunoaffinity chromatography, hydroxyapatite chromatography and lectin chromatography. Other known

-33-

purification methods within the scope of this invention utilize immobilized carbohydrates, epidermal growth factor, or complement domains. Moreover, reverse-phase HPLC and chromatography using ligands for the hybrid immunoglobulin are useful for the purification of the hybrid. Desirably, low concentrations (approximately 1-5 mM) of calcium ion may be present during purification. The LHR may preferably be purified in the presence of a protease inhibitor such as PMSF.

The LHR-immunoglobulin hybrid is employed therapeutically to compete with the normal binding of lymphocytes to lymphoid tissue. The hybrid is therefore particularly useful for organ or graft rejection, and for the treatment of patients with inflammations, such as are for example due to rheumatoid arthritis or other autoimmune diseases. The LHR-immunoglobulin hybrid also finds application in the control of lymphoma metastasis, and in treating conditions in which there is an accumulation of lymphocytes.

LHR-immunoglobulin hybrid heterodimers and heterotetramers are employed in the targeting of therapeutic moieties to lymphoid tissues. For example, a hybrid immunoglobulin consisting of one LHR-IgG chain and one CD4-IgG chain can be used to target CD4-IgG to tissues infected by the viruses such as the human immunodeficiency virus (HIV). Because this hybrid binds to endothelial tissue not only in lymph nodes, but in secondary lymphoid organs such as Peyer's patches and in the brain, it may be used for delivery of CD4-IgG across the blood-brain barrier for the treatment of HIV-related dementia. Similarly, a heterotetrameric immunoglobulin having a LHR-ricin-or CD4-ricin-immunoglobulin as described herein is used to deliver a toxin such as ricin to desired tissues.

In this fashion, selection of ligand binding partners with specific affinity for particular tissues clearly enhances the ability to deliver therapeutic agents which are stable, have relatively long half-lives, and are capable of precise tailoring without undue experimentation.

The novel polypeptide is placed into sterile, isotonic formulations together with required cofactors, and optionally are administered by standard means well known in the field. The formulation is preferably liquid, and is ordinarily a physiologic salt solution containing 0.5 - 10 mM calcium, non-phosphate buffer at pH 6.8-7.6, or may be lyophilized powder.

It is envisioned that intravenous delivery, or delivery through catheter or other surgical tubing will be the primary route for therapeutic administration. Alternative routes include tablets and the like, commercially available nebulizers for liquid formulations, and inhalation of lyophilized or aerosolized receptors. Liquid formulations may be utilized after reconstitution from powder formulations.

The novel polypeptide may also be administered via microspheres, liposomes, other microparticulate delivery systems or sustained release formulations placed in certain tissues including blood. Suitable examples of sustained release carriers include semipermeable polymer matrices in the form of shaped articles, e.g. suppositories, or microcapsules. Implantable or microcapsular sustained release matrices include polylactides (U.S. Patent

-34-

3,773,919, EP 58,481) copolymers of L-glutamic acid and gamma ethyl-L-glutamate (U. Sidman *et al.*, 1985, Biopolymers 22(1): 547-556), poly (2-hydroxyethyl-methacrylate) or ethylene vinyl acetate (R. Langer *et al.*, 1981, J. Biomed. Mater. Res. 15: 167-277 and R. Langer, 1982, Chem. Tech. 12: 98-105). Liposomes containing the hybrid immunoglobulin are prepared by well-known methods: DE 3,218,121A; Epstein *et al.* 1985, Proc. Natl. Acad. Sci. USA, 82:3688-3692; Hwang *et al.*, 1980, Proc. Natl. Acad. Sci. USA, 77:4030-4034; EP 52322A; EP 36676A; EP 88046A; EP 143949A; EP 142541A; Japanese patent application 83-11808; U.S. Patents 4,485,045 and 4,544,545; and UP 102,342A. Ordinarily the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. % cholesterol, the selected proportion being adjusted for the optimal rate of the polypeptide leakage.

Sustained release polypeptide preparations are implanted or injected into proximity to the site of inflammation or therapy, for example adjacent to arthritic joints or peripheral lymph nodes.

The dose of the novel polypeptide administered will be dependent upon the properties of the hybrid employed, e.g. its binding activity and *in vivo* plasma half-life, the concentration of the hybrid in the formulation, the administration route for the hybrid, the site and rate of dosage, the clinical tolerance of the patient involved, the pathological condition afflicting the patient and the like, as is well within the skill of the physician.

The polypeptides of this invention may also be administered along with other pharmacologic agents used to treat the conditions listed above, such as antibiotics, anti-inflammatory agents, and anti-tumor agents. It may also be useful to administer the polypeptide along with other therapeutic proteins such as gamma-interferon and other immunomodulators.

In order to facilitate understanding of the following examples certain frequently occurring methods and/or terms will be described.

"Plasmids" are designated by a lower case p preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

In particular, it is preferred that these plasmids have some or all of the following characteristics: (1) possess a minimal number of host-organism sequences; (2) be stable in the desired host; (3) be capable of being present in a high copy number in the desired host; (4) possess a regulatable promoter; and (5) have at least one DNA sequence coding for a selectable trait present on a portion of the plasmid separate from that where the novel DNA sequence will be inserted. Alteration of plasmids to meet the above criteria are easily performed by those of ordinary skill in the art in light of the available literature and the

teachings herein. It is to be understood that additional cloning vectors may now exist or will be discovered which have the above-identified properties and are therefore suitable for use in the present invention and these vectors are also contemplated as being within the scope of this invention.

5 "Digestion" of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors and other requirements were used as would be known to the ordinarily skilled artisan. For analytical purposes, typically 1 μ g of plasmid or DNA fragment is used with about 2 units of enzyme in about 20
10 μ l of buffer solution. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 μ g of DNA are digested with 20 to 250 units of enzyme in a larger volume. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation times of about 1 hour at 37°C are ordinarily used, but may vary in accordance with the supplier's instructions. After digestion the reaction is
15 electrophoresed directly on a polyacrylamide gel to isolate the desired fragment.

Size separation of the cleaved fragments is performed using 8 percent polyacrylamide gel described by Goeddel, D. *et al.*, *Nucleic Acids Res.*, **8**: 4057 (1980).

"Dephosphorylation" refers to the removal of the terminal 5' phosphates by treatment with bacterial alkaline phosphatase (BAP). This procedure prevents the two restriction
20 cleaved ends of a DNA fragment from "circularizing" or forming a closed loop that would impede insertion of another DNA fragment at the restriction site. Procedures and reagents for dephosphorylation are conventional. Maniatis, T. *et al.*, Molecular Cloning pp. 133-134 (1982). Reactions using BAP are carried out in 50mM Tris at 68°C to suppress the activity of any exonucleases which are present in the enzyme preparations. Reactions are run for 1
25 hour. Following the reaction the DNA fragment is gel purified.

"Oligonucleotides" refers to either a single stranded polydeoxynucleotide or two complementary polydeoxynucleotide strands which may be chemically synthesized. Such synthetic oligonucleotides have no 5' phosphate and thus will not ligate to another oligonucleotide without adding a phosphate with an ATP in the presence of a kinase. A
30 synthetic oligonucleotide will ligate to a fragment that has not been dephosphorylated.

"Ligation" refers to the process of forming phosphodiester bonds between two double stranded nucleic acid fragments (Maniatis, T. *et al.*, Id., p. 146). Unless otherwise provided, ligation is accomplished using known buffers and conditions with 10 units of T4 DNA ligase ("ligase") per 0.5 μ g of approximately equimolar amounts of the DNA fragments to be ligated.

35 "Filling" or "blunting" refers to the procedures by which the single stranded end in the cohesive terminus of a restriction enzyme-cleaved nucleic acid is converted to a double strand. This eliminates the cohesive terminus and forms a blunt end. This process is a versatile tool for converting a restriction cut end that may be cohesive with the ends created

-36-

by only one or a few other restriction enzymes into a terminus compatible with any blunt-cutting restriction endonuclease or other filled cohesive terminus. Typically, blunting is accomplished by incubating 2-15 μ g of the target DNA in 10mM $MgCl_2$, 1mM dithiothreitol, 50mM NaCl, 10mM Tris (pH 7.5) buffer at about 37°C in the presence of 8 units of the
5 Klenow fragment of DNA polymerase I and 250 μ M of each of the four deoxynucleoside triphosphates. The incubation generally is terminated after 30 min. phenol and chloroform extraction and ethanol precipitation.

It is presently believed that the three-dimensional structure of the compositions of the present invention is important to their functioning as described herein. Therefore, all related
10 structural analogs which mimic the active structure of those formed by the compositions claimed herein are specifically included within the scope of the present invention.

It is understood that the application of the teachings of the present invention to a specific problem or situation will be within the capabilities of one having ordinary skill in the art in light of the teachings contained herein. Examples of the products of the present
15 invention and representative processes for their isolation, use, and manufacture appear below, but should not be construed to limit the invention.

EXAMPLES

All references in these examples to the "Mel 14" monoclonal antibody or to "Mel 14" refer to a monoclonal antibody directed against a purported murine form of a lymphocyte
20 surface protein, as described by Gallatin, *et al. supra* Nature 303, 30 (1983). The use of Mel 14 is no longer needed to practice this invention, however, due to the provision herein of full sequences for the DNA and amino acids of the LHR.

Example 1. Purification and Cloning of MLHR.

Isolation of a cDNA Clone Encoding the MLHR.

25 MLHR was isolated from detergent-treated mouse spleens by immunoaffinity chromatography using the Mel 14 monoclonal antibody.

In a typical preparation, 300 spleens from ICR female mice (16 weeks old) were minced and then homogenized with a Potter-Elvehjem tissue grinder in 180 ml of 2% Triton X-100 in Dulbecco's PBS containing 1 mM PMSF and 1% aprotinin. Lysis was continued for 30
30 minutes on a shaker at 4 °C. The lysate was centrifuged successively at 2,000 X G for 5 minutes and at 40,000 X G for 30 minutes.

The supernatant was filtered through Nitex screen and then precleared by adsorption with rat serum coupled to cyanogen bromide-activated Sepharose 4B (10 ml of packed gel). The rat serum was diluted 1:10 for coupling with conjugation carried out according to the
35 manufacturer's instructions. The flow through was applied to a 3 ml column of MEL-14 antibody coupled at 0.5 mg per ml to Sepharose 4B. All column buffers contained sodium azide at 0.02%.

-37-

The column was washed with 25 ml of 2% Triton X-100 in PBS followed by 25 ml of 10 mM CHAPS in the same buffer. Antigen was released by addition of 10 ml of 10 mM CHAPS in 100 mM glycine, 200 mM NaCl, pH 3 and neutralized by collection into 1 ml of 1M TRIS HCl, pH 7.6. After the column was washed with 20 mM triethylamine, 200 mM NaCl, pH 11 and re-equilibrated in 10 mM CHAPS in PBS, the neutralized antigen, diluted into 100 ml of the column buffer, was re-applied and the wash and release steps were repeated.

The purified protein was concentrated in a Centricon 30 (Amicon, Inc.) and analyzed by SDS-PAGE (7.5% acrylamide) with the use of silver staining for visualization. A typical purification yielded 30-40 μ g of antigen per 300 mice based upon comparisons with orosomucoid standards.

In data not shown, a polyacrylamide gel of the purified material showed a diffuse band migrating at approximately 90,000 daltons, and a higher molecular weight protein at around 180,000 daltons. The ratio of the 90,000 dalton to the 180,000 dalton component was 10:1 or greater in all of a large series of preparations. The material was visualized by silver staining of a 10% polyacrylamide gel.

Gas phase Edman degradation of the 90,000 dalton band resulted in the identification of a single N-terminal sequence (Fig. 4A), including the very N-terminal amino acid. 38 N-terminal amino acids were identified, with four gaps (X) at positions 1,19,33, and 34. The asparagine (N) at position 22 was inferred from the absence of an amino acid signal at this position combined with the following tyrosine (Y) and threonine (T) residues, resulting in an N-linked glycosylation site consensus sequence (NXT/S).

The 13-sequence residue shown in Fig. 4A above the 38 residue long N-terminus is that previously deduced by Siegelman *et al. supra* using radioactively-labelled amino acid sequencing, which shows a high degree of homology (11 of 13 residues) with the sequence of the LHR determined here.

No ubiquitin sequence was obtained in any of the three sequencing runs that were done with two separate MLHR preparations. Conceivably, this modification was absent in the mouse splenocytes or the N-terminus of the ubiquitin is blocked to Edman degradation in the LHR from this source.

The amino acid sequences of Fig. 2 were compared with known sequences in the Dayhoff protein data base, through use of the algorithm of Lipman, D. *et al.*, Science 227, 1435-1441 (1981).

The residues in Fig. 4A which are underlined between amino acids 7 and 15 were chosen to produce the oligonucleotide probe shown in Fig. 4B. A 32-fold redundant 26-mer oligonucleotide probe was designed from these residues and synthesized on an Applied Biosystems oligonucleotide synthesizer. All of the possible codon redundancies were included in this probe, with the exception of the proline at position 9, where the codon CCC was chosen based upon mammalian codon usage rules.

Screening of a murine spleen cDNA library obtained from dissected mouse spleens with this probe resulted in the isolation of a single hybridizing cDNA clone. Procedurally, 600,000 plaques from an oligo dT-primed lambda gt 10 murine spleen cDNA library produced from mRNA isolated from murine splenocytes with 5 μ g/ml Concanavalin A for 6 hours were
5 plated at 50,000 phage per plate (12 plates) and hybridized with the P³² labeled 32-fold redundant 26-mer oligonucleotide probe shown in Fig. 4B, in 20% formamide, 5XSSC (150 mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5X Denhardt's solution, 10% dextran sulfate, and 20 micrograms/ml denatured, sheared salmon sperm DNA overnight at 42 °C. These parameters are referred to herein as "stringent conditions". The
10 filters were washed in 1X SSC, 0.1% SDS at 42 °C for 2X 30 minutes and autoradiographed at -70 °C overnight. A single duplicate positive clone was rescreened, the EcoRI insert was isolated and inserted into M13 or PUC 118/119 vectors and the nucleotide sequence determined from single stranded templates using sequence-specific primers.

Figure 2 shows the complete DNA sequence of the 2.2 kilobase EcoRI insert contained
15 in this bacteriophage. The longest open reading frame begins with a methionine codon at position 106-108. A Kozak box homology is found surrounding this methionine codon, suggesting that this codon probably functions in initiating protein translation. A protein sequence containing 373 amino acids of approximately 42,200 daltons molecular weight is encoded within this open reading frame. The translated protein shows a sequence from
20 residues 40 to 76 that corresponds exactly with the N-terminal amino acid sequence determined from the isolated MLHR.

This result suggests that the mature N-terminus of the MLHR begins with the tryptophan residue at position 39. However, it is believed that some proteolytic processing of the actual N-terminus of the LHR may have occurred during the isolation of the protein.

25 A hydrophobicity profile of the protein reveals an N-terminally located hydrophobic domain that could function as a signal sequence for insertion into the lumen of the endoplasmic reticulum. The deduced sequence for positions 39 to 333 is predominantly hydrophilic followed by a 22 residue hydrophobic domain, which is characteristic of a stop transfer or membrane anchoring domain.

30 The putative intracellular region at the very C-terminus of the protein is quite short, only 17 residues in length. On the immediate C-terminal side of the predicted membrane-spanning domain are several basic amino acids, a feature typical of junctions between membrane anchors and cytoplasmic domains of cell surface receptors, Yarden *et al.* Nature. A single serine residue, potentially a site for phosphorylation, is present within the putative
35 cytoplasmic domain.

The protein contains ten potential N-linked glycosylation sites, all of which are within the projected extracellular domain. The absence of asparagine at position 60 (residue 22 of the mature protein) in the peptide sequencing analysis confirms glycosylation at this site and

establishes the extracellular orientation of this region. The coding region contains a total of 25 cysteine residues, although 4 of these cysteine residues are located within the putative leader sequence.

Protein Motifs Within the MLHR

5 As shown in Fig. 6, comparison of the deduced MLHR amino acid sequence to other proteins in the Dayhoff protein sequence databank by using the *fastp* program (Lipman, D., and Pearson, W., *Science* 227, 1435-1441, 1985) revealed a number of interesting sequence homologies.

10 Proteins with the highest sequence homology scores are shown with boxes surrounding the regions of greatest sequence homology. The numbers at the beginning of the sequences show the position within the proteins where these homologous sequences are located.

Fig. 6A shows that the N-terminal motif of the LHR (residues 39 to 155) has certain carbohydrate binding protein homologies, as listed (the percentage of homology of these sequences to the MLHR are given in parentheses, and the references indicated are provided after the Examples): Drickamer; the amino acid residues found by Drickamer *et al.* (1), MuLHR; the MLHR sequence, Hu.HepLec (27.8%); human hepatic lectin (2), Barn.Lec (25%); acorn barnacle lectin (3), Ra. HepLec (23.5%); rat hepatic lectin (4), Ch.HepLec (27.5%); chicken hepatic lectin (5), Hu.IgERec (28.6%); human IgE receptor (6), RaHepLec2 (22.6%); rat hepatic lectin 2 (7), Ra.ASGRec (22.6%); rat asialoglycoprotein receptor (8), Ra.IRP (25.6%); rat islet regenerating protein (9), Ra.MBP (26.1%); rat mannose binding protein (10), Ra.MBDA (26.1%); rat mannose binding protein precursor A (11), Ra.KCBP (27%); rat Kupffer cell binding protein (12), FlyLec (23.1%); flesh fly (*Sarcophaga*) lectin (13), and Rab.Surf (20.9%); rabbit lung surfactant (14).

25 As can be seen, Fig. 6A shows that the most N-terminally localized motif of the LHR shows a high degree of homology with a number of calcium-dependent animal lectins, i.e., C-type lectins (1). These include but are not limited to, various hepatic sugar binding proteins from chicken, rat, and human, soluble mannose-binding lectins, a lectin from Kupffer cells, the asialoglycoprotein receptor, a cartilage proteoglycan core protein, pulmonary surfactant apoproteins, and two invertebrate lectins from the flesh fly and acorn barnacle. Although the complement of "invariant" amino acids initially recognized by Drickamer and colleagues, *supra* as common to C-type animal lectins are not completely conserved in the carbohydrate binding domain of the MLHR, the degree of homology at these residues and at other positions is apparent. The known lectins belonging to the C-type family exhibit a range of sugar-binding specificities including oligosaccharides with terminal galactose, N-acetylglucosamine, and mannose (1).

The fact that there are many residues that are found to be invariant in all of these carbohydrate binding proteins, strongly suggests that this region functions as a carbohydrate

binding domain in the MLHR and apparently explains the observed ability of lymphocytes to bind to the specialized endothelium of lymphoid tissue in a sugar- and calcium-dependent manner. In some embodiments, the carbohydrate binding domain of the LHR alone, without any flanking LHR regions, is used in the practice of this invention.

5 The next motif (residues 160-193) that is found almost immediately after the completion of the carbohydrate binding domain shows a high degree of homology to the epidermal growth factor (egf) family. Fig. 6B shows epidermal growth factor (egf) homologies: MLHR; the MLHR sequence, Notch (38.5%); the *Drosophila melanogaster notch* locus (15), S.purp (31.7%); *Strongylocentrotus purpuratus* egf-like protein (16), Pro.Z (34.1%); bovine protein
10 Z (17), Fact.X (34.2%); coagulation factor X (18), Fact.VII (27.3%); coagulation factor VII (19), Fact.IX (33.3%); coagulation factor IX (20), Lin-12 (32.1%); *Caenorhabditis elegans* Lin-12 locus (21), Fact. XII (26%); coagulation factor XII (22), and Mu.egf (30%); murine egf (23).

As can be seen in Fig. 6B, the greatest degree of homology in this region of the MLHR
15 is found with the *Drosophila* neurogenic locus, *notch*, although there is also significant homology to a number of other members of this large family. The variable location of this domain among the members of this family suggests that this region may be contained within a genomic segment that can be shuffled between different proteins for different functions.

In addition to 6 cysteine residues, virtually all members of this family share three
20 glycine residues. The conservation of cysteine and glycine residues is consistent with the possibility of a structural role for this region in the LHR. It is believed that this domain may place the N-terminally localized carbohydrate binding region in an appropriate orientation for ligand interaction. It is further believed that this domain may serve to strengthen the interaction between the lymphocyte and endothelium by binding to an egf-receptor
25 homologue on the endothelium surface.

The final protein motif in the extracellular region of the MLHR is encoded from amino acids 197 to 328. This region of the glycoprotein contains two direct repeats of a 62 residue sequence that contains an amino acid motif that bears a high degree of homology to a number of complement factor binding proteins (Fig. 6C).

30 Fig. 6C shows complement binding protein homologies: MLHR; MLHR sequence, HuComH (31.9%); human complement protein H precursor (24), MuComH (28.9%); murine complement protein H precursor (25), HuBeta (25.6%); human beta-2-glycoprotein I (26), HuCR1 (29.9%); human CR1 (27), EBV/3d (25%); human Epstein-Barr virus/C3d receptor (28), HuC2 (27.1%); human complement C2 precursor (29), HuB (23.1%); human
35 complement factor B (30), MuC4b (22%); murine C4b-binding precursor (31), HuC1s (29.2%); human C1s zymogen (32), HuC4b (26.1%); human C4b binding protein (33), HuDAF (27.1%); human decay accelerating factor (34), VacSecP (26.2%); vaccinia virus secretory peptide (35).

These proteins, which encode a wide range of multiples of this repeated domain, include, among others, the human and murine complement H precursors, the human beta 2 glycoprotein, the Epstein Barr virus/C3d receptor, the human C4b binding protein, the decay accelerating factor, and the vaccinia virus secretory polypeptide.

5 Figure 7C shows the homologies between the two direct repeats in the MLHR and the direct repeats found in proteins contained within the complement binding family. Many of the amino acids that are conserved in this class of complement binding proteins, including a number of conserved cysteine residues, are also found in the 2 repeats in this region of the MLHR.

10 Interestingly, the two repeats contained within the MLHR are not only exact duplications of each other at the amino acid level, they also show exact homology at the nucleotide sequence level (nucleotide residues 685-865 and 866-1056). While it is possible that this result is due to a cloning artifact, a duplicated region has been found in a number of other clones isolated from a separate cDNA library produced from the MLHR expressing
15 cell line, 38C13 (available from Stanford University, Palo Alto, California, U.S.A.), as well as in a human homologue of the MLHR (discussed, *infra.*). Furthermore, a number of other genes, most notably the Lp(a) gene, show an even higher degree of intragenic repeat sequence conservation of this domain. These results suggest that the MLHR, like other members of the complement binding family, contains multiple repeats of this binding domain.

20 In conclusion, it appears that the extracellular region of the MLHR contains three separate protein motifs that have been joined together to serve a new function or functions. A summary of the protein motifs contained within this glycoprotein is shown in Fig. 7.

Example 2. Cloning of HuLHR.

Generally as described in the previous example, the 2.2 kb EcoR1 insert of the murine
25 Mel 14 antigen cDNA clone described above was isolated, labeled to high specific activity by randomly primed DNA polymerase synthesis with P^{32} triphosphates, and used to screen 600,000 clones from an oligo dT primed lambda gt10 cDNA library derived from human peripheral blood lymphocyte mRNA obtained from primary cells. The filters were hybridized overnight at 42 °C in 40% formamide, 5x SSC (1x SSC is 30 mM NaCl, 3 mM trisodium
30 citrate), 50 mM sodium phosphate (pH6.8), 10% dextran sulfate, 5x Denhardt's solution and 20 micrograms/ml sheared, boiled salmon sperm DNA. They were washed 2x 40 minutes in 0.2x SSC, 0.1% sodium dodecyl sulfate at 55 °C. 12 clones (approximately 1 positive per plate of 50,000 phage) were picked, and the largest EcoR1 insert (~2.2kilobases) was isolated and the DNA sequence was determined by dideoxynucleotide sequencing in the
35 bacteriophage m13 using sequence-specific primers.

This ~2.2 kb clone encoded an open reading frame of 372 amino acids with a molecular weight of approximately 42,200 daltons that began with a methionine which was preceded by a Kozak box homology. The encoded protein contained 26 cysteine residues

-42-

and 8 potential N-linked glycosylation sites. A highly hydrophobic region at the N-terminus of the protein (residues 20-33) was a potential signal sequence, while another highly hydrophobic C-terminally located region of 22 amino acids in length (residues 335-357) was a potential stop transfer or membrane anchoring domain. This C-terminal hydrophobic region was followed by a charged, presumably cytoplasmic, region.

Comparison of the nucleotide sequence of this human clone with that previously found for the MLHR showed a high degree of overall DNA sequence homology (~83%). The relative degrees of amino acid sequence conservation between the MLHR and the HuLHR in each of the LHR domains are: carbohydrate binding domain--83%; egf-like domain--82%; complement binding repeat 1--79%; complement binding repeat 2--63%; overall complement binding domain--71%; and transmembrane domain-- 96%.

Comparison of the published Hermes sequence, Jalkanen, *supra* with the HuLHR sequence of Fig. 1 reveals a lack of sequence homology.

Example 3. Expression of the MLHR.

In order to prove conclusively that the murine cDNA clone isolated here encoded the MLHR, the clone was inserted into an expression vector and analyzed in a transient cell transfection assay. Expression of the HuLHR was performed in a similar fashion.

The Eco R1 fragment containing the open reading frame described above (the ~2.2 kilobase EcoR1 fragment whose sequence is shown in Fig. 2) was isolated and ligated into the pRK5 vector which contains a cytomegalovirus promoter (Eaton, D., *et al.* Biochemistry 25, 8343-8347, 1986; European Publication 307 247 published 15 March 1989). A plasmid containing the inserted cDNA in the correct orientation relative to the promoter was selected and transfected onto 293 human embryonic kidney cells using CaPO₄ precipitation.

After 2 days the cells were incubated with 500 microcuries each of S³⁵ cysteine and methionine. Lysates and supernatants were prepared as previously described (Lasky, L., *et al.* Cell 50, 975-985, 1987) and immunoprecipitated with Mel 14 monoclonal antibody (purified by immunoaffinity chromatography) by utilizing an anti-rat IgG polyclonal antibody in a sandwich between the Mel 14 monoclonal antibody and protein A sepharose.

At the same time, the B-cell lymphoma, 38C13, a cell known to express the MLHR, were either labeled metabolically with either methionine or cysteine, for analysis of the supernatant MLHR, or the cell-surface glycoproteins were labeled with I¹²⁵ and lactoperoxidase for analysis of cell-associated LHR and analyzed by Mel 14 antibody immunoprecipitation.

The resultant immunoprecipitates were analyzed on 7.5% polyacrylamide SDS gels and autoradiographed overnight at -70 °C.

The results of these assays are shown in Fig. 5. In that figure, the lanes A - F signify the following:

-43-

--A. Lysates of 293 cells transfected with a MLHR expression plasmid immunoprecipitated with Mel 14 monoclonal antibody.

--B. Supernatants of 293 cells transfected with a MLHR expression plasmid immunoprecipitated with Mel 14 monoclonal antibody.

5 --C. Lysates of 293 cells transfected with a plasmid expressing the HIV gp120 envelope glycoprotein immunoprecipitated with the Mel 14 monoclonal antibody.

--D. Supernatants of 293 cells transfected with the HIV envelope expression plasmid immunoprecipitated with the Mel 14 monoclonal antibody.

10 --E. Supernatants of 38C13 cells immunoprecipitated with the Mel 14 monoclonal antibody.

--F. Lysates of 38C13 cells surface labeled with I^{125} and immunoprecipitated with the Mel 14 monoclonal antibody.

As can be seen in Fig. 5, cells transfected with this construct produce two cell-associated proteins that reacted specifically with the Mel 14 antibody. The cell associated proteins migrated at approximately ~70,000 daltons and ~85,000 daltons, suggesting that the ~42,200 dalton core protein becomes glycosylated in the transfected cells. The larger band was shifted in molecular weight following sialidase treatment (data not shown), suggesting that it is a relatively mature form of the glycoprotein, whereas the lower molecular weight band was resistant to the enzyme, indicating that it may be a precursor form.

15

20

FACs analysis of transiently transfected cell lines with the Mel 14 antibody showed that a portion of the LHR expressed in these cells was detectable on the cell surface (data not shown).

The higher molecular weight glycoprotein produced in the transfected cell line was found to be slightly smaller than that produced by the Peripheral Lymph Node-homing B-cell lymphoma, 38C13 (Fig. 5, lane F), a result that has been found in other transfected cell lines and may be due to cell-specific differences in glycosylation.

25

Interestingly, both the 38C13 cells and the transfected human cells appeared to shed a smaller molecular weight form of the MLHR into the medium (Fig. 5, lanes B and E). The nature of this shed molecule is unclear, although its reduced molecular weight suggests that it may be a cleavage product of the cell surface form resulting from proteolysis near the membrane anchor.

30

In conclusion, these results convincingly demonstrate that the cDNA clone that we have isolated encodes the MLHR.

35 **EXAMPLE 4. Construction, Purification, and Analysis of**

Truncated MLHR-IgG Chimeras.

Fig. 8 shows the construction of MLHR-IgG chimeras containing the lectin, lectin-egf, and lectin-egf-complement regulatory motifs. The top of the figure illustrates the protein

-44-

domains of the murine lymphocyte homing receptor (MLHR) including the N-terminal signal sequence (SS), the lectin, epidermal growth factor (egf), and duplicated complement regulatory domains (CDB) as well as a transmembrane anchor domain (TMD) and a short cytoplasmic sequence. The three truncated MLHR-IgG chimeras which contain the lectin (MLHR-L + IgG), the lectin and egf (MLHR-LE + IgG) and the lectin, egf, and two complement regulatory motifs (MLHR-LEC + IgG) are also shown in Fig. 8. These truncated proteins are all joined to a human heavy chain gamma 1 region just upstream of the hinge domain (H) such that these chimeras contain the two cysteine residues (C) of the hinge responsible for immunoglobulin dimerization as well as the CH2 and CH3 constant regions. A previously characterized human heavy chain IgG 1 constant region cassette (Caponet *et al.*, *supra* 1989) was utilized. Junctional sites between the LHR and human IgG sequences was chosen such that the joining of the molecules near the hinge region resulted in chimeric molecules that were efficiently synthesized and dimerized in the absence of any light chain production. In addition, the use of the human IgG 1 constant region obviates any difficulties due to cross reactivity with endogenous murine IgGs in the immunohistochemical experiments described below.

As can be seen from figure 9, these chimeras were efficiently synthesized and secreted in these transient transfection assays. The reactivity of these chimeras with protein A sepharose in the absence of added antibodies demonstrates that the constant region domains are normally folded. Figure 9 illustrates that these molecules dimerize under non-reducing conditions, demonstrating that the hinge region is fully functional in these chimeras. Finally, the protein A reactivity also allows for the purification of these chimeras to near homogeneity on protein A sepharose columns. The results herein demonstrate the production of antibody-like entities whose "variable" domain may be said to be derived from the MLHR while the constant domain is derived from the human IgG gamma 1 heavy chain.

Construction of Chimeras

Starting with a previously described MLHR-PRK5 expression plasmid (Eaton *et al.*, 1986,; Lasky *et al.*, *Cell* 50:975-985, 1987) and a cDNA copy of a human heavy chain IgG (Capon *et al.*, *Nature* 337:525-531, 1989), an 1100 bp HindIII fragment encoding the CH1-CH3 regions of the human IgG 1 constant region was inserted 3 prime of the polyA site of the MLHR cDNA. This plasmid was converted to single stranded template by utilizing an m13 origin of replication and the K07 helper phage, after which regions between the hinge and the lectin, egf, and second complement binding repeat N-terminal to the putative trans membrane anchoring region) were looped out with 48-mer oligonucleotides by *in vitro* mutagenesis (Zoller and Smith, 1982). The resultant mutants were screened with 32P-labeled 21-mer oligonucleotides spanning the deletion junctions, and the isolated mutants were sequenced using supercoil sequencing.

Correct mutants were tested for expression by transfection onto human kidney 293 cells using previously described methods. 35S methionine and cysteine labeled supernatants were analyzed by immunoprecipitation with protein A sepharose beads in the absence of added antibodies. The precipitated proteins were analyzed on 7.5% polyacrylamide-SDS gels either with or without reduction with beta mercaptoethanol. Plasmids that resulted in correctly expressed chimeras were introduced into 293 cells by transfection in the presence of selective plasmids encoding resistance to G418 as well as dihydrofolate reductase. Clones were selected in G418, and the incorporated plasmids were amplified in the presence of methotrexate. Permanent cell lines expressing high levels of each construct were grown to large scale in T-flasks, and the cell supernatants were clarified by centrifugation and filtration. The resultant supernatants were concentrated by Amicon filtration and passed over standard protein A-sepharose columns, washed with PBS, and eluted with 0.1M Acetic Acid, 0.15 M NaCl (pH 3.5). The eluted material was immediately neutralized with 3 M Tris, pH 9, and quantitated by SDS gel electrophoresis as well as an ELISA assay.

The gel electrophoresis results described in the preceding paragraph are shown in Fig. 9. Reduced Proteins are shown in lanes A-F., Non-reduced proteins in lanes G-I, and purified proteins in lanes J-L. Molecular weights of markers are shown in kilodaltons. Lane identifications are as follows: A. Secreted MLHRLEC-IgG, B. Intracellular MLHRLEC-IgG, C. Secreted MLHRLE-IgG, D. Intracellular MLHRLE-IgG, E. Secreted MLHRL-IgG, F. Intracellular MLHRL-IgG., G. Secreted MLHRLEC-IgG, H. Secreted MLHRLE-IgG, I. Secreted MLHRL-IgG, J. Purified MLHRLEC-IgG, K. Purified MLHRLE-IgG, and L. Purified MLHRL-IgG.

Isolated LHR-IgG Chimeras were quantitated using an ELISA format that consisted of an anti-human IgG 1-specific mouse monoclonal antibody coating microtitre wells. Unknown samples as well as highly purified human CD4-IgG 1 immunoadhesin standard were incubated with antibody-coated plates, after which the plates were washed, and the bound material was reacted with horse radish peroxidase-conjugated goat-anti human IgG 1, followed by further washes and addition of substrate. This quantitative assay allowed for the measurement of sub-nanogram quantities of isolated LHR-IgG chimeras.

Analysis of MLHR-IgG Chimera PPME Reactivity by ELISA.

The ability of various IgG chimeras to recognize the yeast cell wall carbohydrate, polyphosphomannan ester or PPME, was analyzed in an ELISA format as previously described (Imai *et al.* 1989). Briefly, approximately equivalent amounts of the purified chimeras were coated onto microtitre wells overnight at 4 °C. Non-specific sites were blocked with BSA, after which the bound antigens were reacted with a 5 microgram per ml solution of PPME. Bound carbohydrate was detected with a polyclonal antibody directed against it and standard (Vector) immunohistochemical staining reagents. Inhibition with Mel 14 was performed by pre-incubating MLHRLEC-IgG containing wells with the monoclonal antibody before the addition of PPME, while the calcium dependance of the homing receptor-carbohydrate

interaction was demonstrated by inclusion of 10 mM EGTA during the binding reaction. Various other additives were added before PPME incubation in assays examining inhibition. After 1 hr at 22 °C, the plates were washed and incubated with a rabbit polyclonal antibody directed against PPME for 1 hr at 22 °C. Plates were washed and incubated with Vector ABC-AP for 30 minutes, washed, and developed. The resulting assays were measured on a plate reader. Carbohydrates used in the inhibitory assays were obtained from Sigma Chemical Co. (St. Louis, MO.)

Results of the PPME binding analysis are shown in Fig. 10. The lanes contain the following MLHR-IgG chimeras: A. Binding of PPME to MLHRL-, MLHRLE- and MLHRLEC-IgG chimeras. B. Inhibition of MLHRLEC-IgG-PPME binding with Mel 14 monoclonal antibody and EGTA. C. Inhibition of MLHRLEC-IgG-PPME binding with other carbohydrates.

Previous work had demonstrated that the LHR was able to bind to a yeast cell wall mannan, polyphosphomannan ester or PPME (Yednock *et al.*, J. Cell Biol. 104:725-731, 1987), and that this binding inhibited the ability of lymphocytes to adhere to peripheral lymph node high endothelial vesicles, in agreement with the supposition that the peripheral lymph node LHR lectin domain may recognize a carbohydrate on the peripheral lymph node endothelium. In addition, the MEL 14 antibody was found to inhibit the binding of PPME to the lymphocyte surface (Yednock *et al.*, *supra* 1987), consistent with the notion that this carbohydrate bound within the lectin domain of the peripheral lymph node LHR.

The chimera that contained the lectin, egf, and duplicated complement binding repeat structures was found to bind PPME. This binding was inhibitable by the Mel 14 antibody, in agreement with data demonstrating that the MLHRLEC-IgG chimera was recognized by this antibody (data not shown), and was quantitatively comparable to that found previously using MLHR isolated from spleen cells (Imai *et al.*, submitted for publication, 1989), suggesting that it represented the same protein-carbohydrate interaction that has been found with the LHR on the lymphocyte surface (Yednock *et al.*, *supra* 1987). In addition, the binding was also found to be calcium dependent (Stoolman and Rosen, J. Cell Biol. 96:722-729, 1983), implying that the type C or calcium-dependent lectin domain (Drickamer, J. Biol. Chem. 263:9557-9560, 1988) was at least partly responsible for this interaction, as has been shown for the lymphocyte-associated receptor (figure 9b).

Previous work demonstrated that a variety of carbohydrates besides PPME were capable of being recognized by the spleen derived MLHR (Yednock *et al.*, *supra* 1987; Imai *et al.*, *supra* 1989). These included fucoidin, dextran sulfate, and brain derived sulfatides. The ability of these carbohydrates to inhibit the interaction between the MLHRLEC-IgG chimera and PPME was examined to investigate the specificity of this molecule versus the previously described spleen-derived glycoprotein (Imai *et al.*, *supra* 1989). As can be seen from figure 9, fucoidin, dextran sulfate, and sulfatide are all able to inhibit the interactions between PPME and MLHRLEC-IgG, implying that the carbohydrate specificity of this

recombinant-derived protein mimics that previously described for the naturally occurring protein. The lack of inhibition by two other charged carbohydrates, chondroitin sulfate and heparin, suggests that the inhibition is due to specific carbohydrate recognition and not to non-specific interference due to the highly charged nature of the compounds.

5 **Cell Blocking Assays with MLHR-IgG Chimeras.**

The Stampfer-Woodruff cell blocking assay (Stampfer and Woodruff, J. Exp. Med. 144:828-833, 1976) was performed with cryostat-cut sections of mouse peripheral lymph nodes as well as with Peyer's patch as previously described (Geoffrey and Rosen, J. Cell Biol., in press, 1989). Briefly, the frozen tissue sections were incubated with mesenteric
10 lymphocytes in the presence of either the MLHR-IgG chimeras, isolated spleen-derived MLHR, or buffer alone. MLHR-IgG chimeras were included at concentrations as high as 10 micrograms per section and were pre-incubated on frozen sections before the addition of 1×10^7 cells per ml. The slides were washed, and lymphocyte attachment was measured by digital morphometry as the number of lymphocytes bound to HEV in these lymphoid organs
15 per unit area.

In data not shown, the MLHRLEC-IgG chimera was found to inhibit the binding of lymphocytes to peripheral lymph node HEV at a level of approximately 75% inhibition while, in this assay, the spleen-derived MLHR blocked at a level of about 50%. This inhibition was calcium dependent and was blocked by the inclusion of the MEL 14 monoclonal antibody
20 (data not shown).

Immunohistochemical Analysis of MLHR-IgG Chimeras.

Isolated MLHR-IgG chimeras were utilized for immunohistochemical experiments using procedures identical to those used for monoclonal antibodies. 8-10 micron tissue sections were cut in a cryostat and fixed with 0.1 M cacodylate, 1% paraformaldehyde for 30 minutes
25 at 4 °C. The sections were washed in Dulbecco's PBS and stained with varying amounts of MLHR-IgG chimera in 5% normal mouse serum at 4 °C for 30 minutes. The sections were then washed and incubated with a second stage containing biotinylated goat anti-human Fc specific antibody (Vector). Endogenous peroxidase was eliminated by treating the sections with hydrogen peroxide-methanol after the addition of the second stage reagent and before
30 the addition of the Vector ABC complex. Sections were washed and incubated with substrate (AEC) for 5-10 minutes. Finally, the sections were counter-stained with aqueous hematoxylin (Biomedica) and viewed with a Zeiss Axioplot.

These immunohistochemical analyses of the three MLHR-IgG chimeras used peripheral lymph node as a tissue source. The choice of peripheral lymph node as a histology source
35 was dictated by the large body of previous literature which demonstrated that lymphocytes bind to the HEV of this lymphoid tissue in a manner which can be blocked by Mel-14, implying that the highest level of ligand recognized by the MLHR should be in this tissue (Gallatin *et al.*, Nature 304:30-34, 1983). The MLHRLEC-IgG chimera was able to stain

-48-

peripheral lymph node HEV. The staining was found exclusively over the high walled endothelial cells, with no staining of ad- or abluminal regions. In addition, this staining could be blocked by the MEL 14 antibody and was dependent upon the presence of calcium, suggesting that the binding of MLHRLEC-IgG to peripheral lymph node HEV mimicked the adhesion between lymphocytes and the HEV. In concordance with the PPME binding data, the staining of peripheral lymph node HEV by MLHRLEC-IgG was inhibitable by fucoidin and dextran sulfate (figure 5), while chondroitin sulfate and simple mannans were incapable of inhibiting the staining reaction (data not shown), again implying that the staining reaction was due to the recognition of a carbohydrate ligand expressed on the peripheral lymph node HEV. These data reveal that this type of immunohistochemical reagent may be utilized to investigate the tissue distribution of the endothelial molecule(s) which are capable of interacting with the peripheral lymph node LHR.

The MLHR Ligand is Found in Peyer's Patches

We have found, in results of immunohistochemical assays not shown, that the MLHRLEC-IgG chimera is, surprisingly, able to recognize the endothelium of Peyer's patches specifically. The chimera appears to stain the high walled endothelium of Peyer's patches vessels containing lymphocytes. This staining is inhibitable by the MEL 14 antibody and is also calcium dependent. Interestingly, the staining of the Peyer's patches HEV appears somewhat weaker relative to that found for the staining of the peripheral lymph node HEV, implying that a lower level of the MLHR ligand(s) may be expressed in this lymphoid organ. These results demonstrate that, while other adhesion systems may be involved in this organ (Holzman *et al.*, *Cell* 56:37-46, 1989), the ligand(s) for the peripheral lymph node LHR is expressed and, therefore, is involved in lymphocyte binding to the endothelium of this lymphoid organ.

EXAMPLE 5. Construction of CD4-IgG-MLHR-IgG Chimeras.

Two previously constructed PRK plasmids were used to direct expression of MLHR-IgG and human CD4-IgG. The MLHR plasmid is as described in the previous example. The CD4-Ig plasmid is that described in Capon *et al.*, *supra*, modified by the deletion of the coding region for the C_{H1} domain and a portion of the hinge region up to the first cysteine residue. These plasmids were cotransfected by the standard calcium-phosphate method as described above into human 293 cells, either together with PSV^{T antigen} to generate cells transiently expressing the two genes at high levels, or together with PSV^{neo}, to confer neomycin resistance for selection of cell clones stably expressing the two genes. Expression was analyzed by radioimmunoprecipitation; because CD4-IgG, LHR-IgG and CD4-IgG-LHR-IgG all contain an IgG Fc portion, they can all be precipitated directly by protein A by standard methods. Three types of molecules were detected: CD4-IgG homodimers, LHR-IgG homodimers, and CD4-IgG-LHR-IgG heterodimers. These molecules are separated to their

-49-

monomeric constituents by reduction, indicating that the members of each dimer, including heterodimers, are covalently attached to one another by disulfide bonds.

REFERENCES TO THE EXAMPLES

1. Drickamer, K., J. Biol. Chem. 263, 9557 (1988); Drickamer, K., Kidney Int. 32, S167 (1987).
2. Spiess, M., *et al.*, Proc. Natl. Acad. Sci., U.S.A. 82, 6465 (1985).
- 5 3. Muramoto, K., *et al.*, Biochem Biophys. Acta 874, 285 (1986).
4. Leung, J., *et al.*, J. Biol. Chem. 260, 12523 (1985); Holland, E., *et al.*, Proc. Natl. Acad. Sci., U.S.A. 81, 7338 (1984).
5. Drickamer, K., J. Biol. Chem. 256, 5827 (1981).
6. Kikutani, H., *et al.*, Cell 47, 657 (1986).
- 10 7. McPhaul, M., *et al.*, Molec. Cell. Biol. 7, 1841 (1987).
8. Halberg, D., *et al.*, J. Biol. Chem. 262, 9828 (1987).
9. Terzaono, K., *et al.*, J. Biol. Chem. 263, 2111 (1988).
10. Drickamer, K., *et al.*, J. Biol. Chem. 262, 2582 (1987).
11. Drickamer, K., *et al.*, J. Biol. Chem. 261, 6878 (1986).
- 15 12. Hoyle, G., *et al.*, J. Biol. Chem. 263, 7487 (1988).
13. Takahashi, H., *et al.*, J. Biol. Chem. 260, 12228 (1985).
14. Boggaram, V., *et al.*, J. Biol. Chem. 263, 2939 (1988).
15. Kidd, S., *et al.*, Mol. Cell. Biol. 6, 3094 (1986).
16. Hursh, D., *et al.*, Science 237, 1487 (1987).
- 20 17. Hojrup, P., *et al.*, FEBS Lett. 184, 333 (1985).
18. Fung, M., *et al.*, Nucl. Acids Res. 12, 4481 (1984).
19. Takeya, H. *et al.*, Proc. Natl. Acad. Sci., U.S.A. 76, 4990 (1979).
20. McMullen, B., *et al.*, Biochem Biophys. Res. Commun. 115, 8 (1983).
21. Greenwald, I., Cell 43, 583 (1985).
- 25 22. Cool, D., *et al.*, Biol. Chem. 260, 13666 (1985).
23. Gray, A., *et al.*, Nature 303, 722 (1983).
24. Schulz, T., *et al.*, Eur. J. Immunol. 16, 1351 (1986).
25. Kristensen, T., *et al.*, Proc. Natl. Acad. Sci., U.S.A. 83, 3963 (1986).
26. Lozier, J., *et al.*, Proc. Natl. Acad. Sci., U.S.A. 81, 3640 (1984).
- 30 27. Bentley, D. Biochem. J. 239, 339 (1986).
28. Moore, M., *et al.*, Proc. Natl. Acad. Sci., U.S.A. 84, 9194 (1987).
29. Bentley, D., *et al.*, Proc. Natl. Acad. Sci., U.S.A. 81, 1212 (1984).
30. Mole, J., *et al.*, J. Biol. Chem. 259, 3407 (1984).
31. DiScipio, R., *et al.*, J. Biol. Chem. 263, 549 (1988).
- 35 32. Kusomoto, H., *et al.*, Proc. Natl. Acad. Sci., U.S.A. 85, 7307 (1988).
33. Lintin, S., *et al.*, FEBS Lett. 204, 77 (1986).
34. Caras, I., *et al.*, Nature 325, 545 (1987).
35. Kotwal, G., *et al.*, Nature 335, 176 (1988).

Claims:

1. A polypeptide fusion comprising a ligand binding partner protein and a stable plasma protein, wherein the ligand binding partner protein is not a platelet growth factor receptor or an insulin receptor.
- 5 2. The polypeptide of claim 1, wherein the stable plasma protein is an immunoglobulin chain and the ligand binding partner protein and immunoglobulin are fused through C- or N-terminal amino or carboxyl groups.
3. The polypeptide of claim 2, wherein a portion of the constant domain of the immunoglobulin is linked at its N-terminus to the C-terminus of said ligand binding partner.
- 10 4. The polypeptide of claim 2, wherein the Fc portion of the constant domain of the immunoglobulin is linked at its N-terminus to the C-terminus of said ligand binding partner.
5. The polypeptide of claim 2, wherein the immunoglobulin sequence is obtained from IgG1, IgG2, IgG3, IgG4, IgA, IgE, IgD or IgM.
6. The polypeptide of claim 2, wherein the ligand binding partner comprises a LHR.
- 15 7. The polypeptide of claim 2, wherein the ligand binding partner comprises a cell surface adhesion molecule.
8. The polypeptide of claim 2, joined to a detectable marker.
9. The polypeptide of claim 2, wherein said immunoglobulin sequence is human.
10. The polypeptide of claim 2, wherein said ligand binding partner is human.
- 20 11. The polypeptide of claim 2, wherein the constant domain is the constant domain of an IgG heavy chain.
12. The polypeptide of claim 2, wherein the ligand binding partner is an LHR which binds to the high endothelial venules of lymphoid tissue.
13. The polypeptide of claim 2, unassociated with glycosylation.
- 25 14. The polypeptide of claim 2 in a physiologically acceptable carrier.
15. The polypeptide of claim 14, wherein the carrier is a sterile, isotonic solution.
16. The polypeptide of claim 14, wherein the carrier is a sustained-release formulation.
17. The polypeptide of claim 14, wherein the carrier is a liposome.
- 30 18. The polypeptide of claim 2, wherein said ligand binding partner protein in the fusion is incapable of cell membrane anchorage.
19. The polypeptide of claim 2, wherein said immunoglobulin is from a different species than said binding partner.
20. The polypeptide of claim 2, wherein the ligand binding partner is a single chain.
- 35 21. The polypeptide of claim 2, wherein the ligand binding partner contains more than one polypeptide chain, one chain of which is fused to an immunoglobulin constant region.

22. The polypeptide of claim 21, wherein the fused chain of the ligand binding partner is devoid of its transmembrane and cytoplasmic domains.

23. The polypeptide of claim 2, wherein said ligand binding partner is a cell membrane protein.

5 24. The polypeptide of claim 2, wherein said ligand binding partner is a constant region-like domain of an immunoglobulin superfamily member.

25. The polypeptide fusion of claim 2, wherein the ligand binding partner is a first ligand binding partner, said fusion further comprising an additional fusion of a second ligand binding partner and an immunoglobulin.

10 26. The polypeptide fusion of claim 25 wherein the second ligand binding partner is different from the first ligand binding partner.

27. The polypeptide fusion of claim 26 wherein the first ligand binding partner is LHR and the second ligand binding partner is CD4.

15 28. The polypeptide fusion of claim 25 wherein the first ligand binding partner fusion and second ligand binding partner fusion are crosslinked by disulfide bonds.

29. The polypeptide fusion of claim 25 wherein the first ligand binding partner fusion and second ligand binding partner fusion are noncovalently aggregated.

20 30. The polypeptide of claim 27, wherein said ligand binding partner is a LHR having a LHR carbohydrate binding domain, a LHR epidermal growth factor domain, and a LHR complement binding domain, and wherein said LHR carbohydrate binding domain has been deleted or replaced by a heterologous carbohydrate binding domain.

25 31. The polypeptide of claim 27, wherein said ligand binding partner is a LHR having a LHR carbohydrate binding domain, a LHR epidermal growth factor domain, and a LHR complement binding domain, and wherein said LHR epidermal growth factor binding domain has been deleted or replaced by a heterologous epidermal growth factor binding domain.

32. The polypeptide of claim 27, wherein said ligand binding partner is a LHR having a LHR carbohydrate binding domain, a LHR epidermal growth factor domain, and a LHR complement binding domain, and wherein said LHR complement binding domain has been deleted or replaced by a heterologous complement binding domain.

30 33. The polypeptide of claim 27, wherein said ligand binding partner is a LHR having a LHR carbohydrate binding domain, a LHR epidermal growth factor domain, a LHR complement binding domain, a LHR cytoplasmic domain and a LHR transmembrane domain, and wherein said LHR cytoplasmic and transmembrane domains have been inactivated.

35 34. The polypeptide of claim 1 wherein the stable plasma protein is selected from the group consisting of serum albumin, transferrin, lipoprotein and apolipoprotein.

35. The polypeptide of claim 34 wherein the plasma protein is albumin.

36. The polypeptide of claim 1 wherein the ligand binding partner is a CD antigen.

37. The polypeptide of claim 1 wherein the ligand binding partner is a LHR.

38. A polypeptide fusion comprising a first ligand binding partner and a stable plasma protein, said fusion further comprising an additional fusion of a second ligand binding partner and a second stable plasma protein, wherein said first and second ligand binding partners are different.

5 39. A polypeptide fusion comprising a ligand binding partner and a stable plasma protein, wherein said stable plasma protein is not an immunoglobulin.

40. A polypeptide fusion comprising a ligand binding partner and a stable plasma protein, said fusion comprising a single chain fusion.

10 41. A polypeptide fusion comprising a first ligand binding partner and a stable plasma protein, said fusion further comprising an additional fusion of a second ligand binding partner and a second stable plasma protein, wherein each of said stable plasma proteins are immunoglobulin constant domains.

42. Nucleic acid encoding the polypeptide of claim 1.

43. A replicable expression vector comprising the nucleic acid of claim 42.

15 44. A composition comprising a cell transformed with the recombinant expression vector of claim 43.

45. The composition of claim 44 wherein the cell is a mammalian cell.

46. The composition of claim 44 wherein the cell is a chinese hamster ovary cell line.

20 47. A method of culturing the cell of claim 44 which comprises culturing the transformed cell and recovering a polypeptide from the cell culture.

48. The method of claim 47 wherein the polypeptide is recovered from the host cell.

49. The method of claim 47 wherein the polypeptide is secreted into the culture medium and recovered from the culture medium.

GAATTCAGATGTCGCTGCTTCTCACTGCA.....ACACAGCACTCCCTTTGCCAAGCACTGACACCTTTGTCTAGCT

1
met ile phe pro trp lys [Cys
ATG ATA TTT CCA TCG AAA [TGT]

Signal Sequence
20
gln ser thr gln arg asp leu trp asn ile phe lys leu trp gly trp thr met leu [Cys
CAG ACC ACC CAG AGG GAC TTA TCG AAC ATC TTC AAG TTT TGG GGG TCG ACA ATG CTC [TGT]

Probable N-Terminus
30
[Cys] asp phe leu ala [his his gly thr tyr [Cys] trp thr tyr his tyr ser gln lys pro
[TGT] GAT TTC CTG GCA [CAT CAT GCA ACC TAC [TGT] TCG ACT TAC CAT TAT TCT GAA AAA CCC

80
met asn trp gln arg ala arg arg phe [Cys] arg asp [asn tyr thr] asp leu val ala ile
ATG AAC TCG CAA AGG GCT AGA AGA TTC [TGT] GCA GAC [AAT TAC ACA] CAT TTA GTT GCC ATA

70
gln asn lys phe gln ile gln tyr leu gln lys thr leu pro phe ser arg ser tyr tyr
CAA AAC AAG GCG GAA ATT GAG TAT CTG GAG AAG ACT CTG CCC TTC AGT GGT TCT TAC TAC

90
trp ile gly ala atc cgg aag ata gga gaa ata tgg acg 100 trp val gly thr [asn lys ser] leu
TGG ATA GCA ATC CCG AAG ATA GGA GAA ATA TGG ACG TCG CTG GCA ACC [AAC AAA TGT] CTC

110
thr gln gln ala gln asn trp gly asp gly gln pro 120 asn lys lys asn lys gln asp
ACT GAA GAA GCA GAG AAC TGG GCA CAT GGT GAG CCC AAC AAC AAG AAG AAC AAG GAG GAC

130
[Cys] val gln ile tyr ile lys arg asn lys asp ala 140 gly lys trp asp asp ala [Cys
[TGT] GTG GAG ATC TAT ATC AAG AGA AAC AAA GAT GCA GGC [TGT] AAC TGG AAC GAT GAG CCC [TGT]

150
his lys leu lys ala ala leu [Cys] tyr thr ala ser 160 gln pro trp ser [Cys] ser gly
CAC AAA CTG AAG GCA GCC CTC [TGT] TAC ACA GCT TCT [TGT] CAG CCC TGG TCA [TGT] AGT GGC

170
his gly gln [Cys] val gln ile ile asn [asn his thr] [Cys] 180 [asn] asp val gly tyr tyr
CAT GCA GAA [TGT] GTA GAA ATC ATC AAT [AAT CAC ACC] [TGT] AAC [TGT] CAT GTG GGG TAC TAT

190
gly pro gln [Cys] gln leu val ile gln [Cys] gln pro leu gln ala pro gln leu gly thr
GGG CCC CAG [TGT] CAG CTG GTG ATT CAG [TGT] GAG CCT TTT GAG GCC CCA GAG CTG GCT ACC

210
met asp [Cys] thr his pro phe gly [asn phe ser] phe ser ser gln [Cys] ala phe ser [Cys
ATG GAC [TGT] ACT CAC CCC TTT GCA [AAC TTC AGC] TTC ACC TCA CAG [TGT] GCC TTC AGC [TGT]

230
ser gln gly thr [asn leu thr] gly ile gln gln thr 240 [Cys] gly pro phe gly [asn trp
TCT GAA GCA ACA [AAC TTA ACT] CGG ATT GAA GAA ACC ACC [TGT] GCA CCA TTT GCA [AAC TGG

250
[ser] ser pro gln pro thr [Cys] gln val ile gln [Cys] 260 pro leu ser ala pro asp leu
TCA TCT CCA GAA GCA ACC [TGT] CAA GTG ATT CAG [TGT] GAG CCT CTA TCA GCA GCA CAT TTT

270
gly ile met [asn [Cys] ser] his pro leu ala ser phe 280 ser phe thr ser ala [Cys] thr phe
GGG ATC ATG [AAC [TGT] ACC] CAT CCC CTG GCC ACC TTC ACC TTT ACC TCT GCA [TGT] ACC TTC

290
ile [Cys] ser gln gly thr gln leu ile gly lys lys 300 lys thr ile [Cys] gln ser ser gly
ATC [TGT] TCA GAA GGA ACT GAG TTA ATT GGG AAG AAG AAA ACC ATT [TGT] GAA TCA TCT GCA

310
ile trp ser [asn pro ser] pro ile [Cys] gln lys leu 320 ser lys ser phe ser met ile lys
ATC TGG TCA [AAT CCT AGT] CCA ATA [TGT] CAA AAA TTT GAG AAA AGT TTC TCA ATG ATT AAG

330
gln gly asp tyr asn pro leu phe ile pro val ala 340 Stop Transfer Sequence
GAG GGT GAT TAT AAC CCC CTC TTC ATT CCA CTG GCA GTT ATG CTG TCT ACT GCA TTC TCT GCG

350
leu ala phe ile ile trp leu ala arg arg leu lys lys gly lys lys ser lys arg ser
TTG GCA TTT ATC ATT TGG CTG GCA AGG AGA TTA AAA AAA GCG AAG AAA TCC AAG AGA AGT

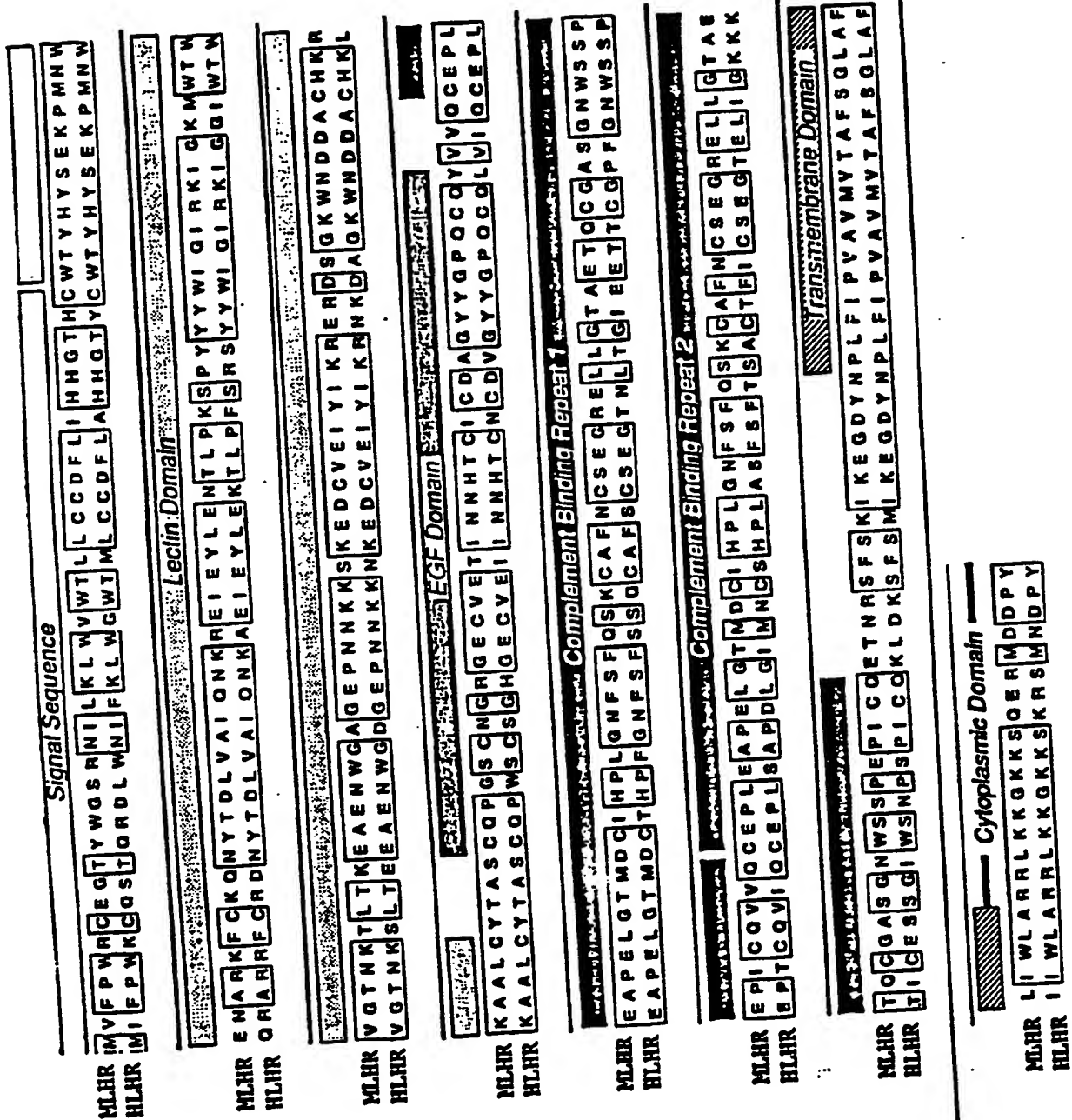
370 372
met asn asp pro tyr GC
ATG AAT GAC CCA TAT TAA ATCGCCCTTGCTGAAGCAAAATTCCTGCAATACTAAAAATCATGAGATCTTTTA

AAATCCTTCATGAAGCTTTTGTGTGGTGGCAGCTCTACGTCAAAATGAAGTGTGTCTTCAGTGCATCTGCGAA
GATTTCTACCCGAGCAACAGTTTCCTTCAGCTTCCATTTCCGCTCCTCAATTATCCCTCAACCCCGAGCCACAGCTGTT
TATACAGCTCAGCTTTTGTCTTTCTTCAGCAGCAAAATAAGACCATAAGCGAAGCAATTCATGTGCAATATAAG
ATGGCTGACTTTGCTCTTTCTTGACTCTTTCTTCAGTTTCATTCAGTCTGTACTTGATCAGCAGCACTTCTAAAT
GAAGTGCAAAATTTGATACATATGTCAATATGCACTCAGTTTCTTGCAGATCAAAATTTACGTTGCTTCTCTGTATACT
GTGAGGTACACTCTTATAGAAAGTTCAAAAAGTCTACGCTCTCTCTTTCTTCTAACTCAAGTCAAGTAAATGGGCTC
TGCTCAAGTTCAAGAGTCTTATTCAGTGTAGCTCTGCGGTCTGTGAATTCGACCACTCATTTAACTGGCTTCAG
GCTTCGCACTTCTTCAGCCACTCTCTTTTCAGTTGGCTGACTTCCACACTACCATCTCATGAGTGCAGCA
AAGCAGACAGCAGCAAAATAGCTCTGCGGTTTTTTTACTTTGGCGTTTTTCTGTTTCTTTTATGACACCACTTCTT
ATTTCTTATAGTCAATGTTTTCTTTATCAGCATATTATTAAGAAAACATCACTGAAATCTAGCTGCAAGTGACA
TCTCTTTCATCTCATGCAAGCTTAAACAGCTGCAAAATTCCTTGATTCAAAATGAATGCTCTCTTTGCTT
GCTCCAGCACTTTTATCACTTACCTAGATTTCTACATATTTCTTTAAATTTCAATCTCAGCCCTCTCTCAACCCGAGG
GCTTCAGCAGCACTGCAATTC

Fig. 1

Fig. 3

3/12



4/12

Fig. 4A

T H K M KF K VVIL K
 10 20 30
(TYHYSEKPMNWENARKFXKQNYTDLVAIQNKXXIEYL

Fig. 4B

5' GAG AAG CCC ATG AAT TGG GAG AAT GC 3'

5/12

Fig. 5



6/12

Fig. 6

Drickamer	Q	C
MLHR	10	10
Hu.HepLec	20	20
Barn.Lec	30	30
Ra.HepLec	40	40
Ch.HepLec	50	50
Hu.IgErec	60	60
Ra.HepLec2	70	70
Ra.ASGRec	80	80
Ra.IRP	90	90
Ra.HBP	100	100
Ra.MBDA	110	110
Ra.KCBP	120	120
FlyLec	130	130
Rab.Surf	140	140

Drickamer	Q	C
MLHR	10	10
Hu.HepLec	20	20
Barn.Lec	30	30
Ra.HepLec	40	40
Ch.HepLec	50	50
Hu.IgErec	60	60
Ra.HepLec2	70	70
Ra.ASGRec	80	80
Ra.IRP	90	90
Ra.HBP	100	100
Ra.MBDA	110	110
Ra.KCBP	120	120
FlyLec	130	130
Rab.Surf	140	140

Drickamer	Q	C
MLHR	10	10
Hu.HepLec	20	20
Barn.Lec	30	30
Ra.HepLec	40	40
Ch.HepLec	50	50
Hu.IgErec	60	60
Ra.HepLec2	70	70
Ra.ASGRec	80	80
Ra.IRP	90	90
Ra.HBP	100	100
Ra.MBDA	110	110
Ra.KCBP	120	120
FlyLec	130	130
Rab.Surf	140	140

7/12

Fig. 6B

MLHR	100	C	Q	P	G	S	C	N	G	R	G	E	C	V	E	T	I	N	N	H	T	C	I	C	D	A	G	Y	Y	G	P	O	C	Q	Y		
Notch	1021	C	T	E	S	S	C	L	N	G	G	S	C	I	D	G	I	N	G	Y	N	C	S	C	L	A	G	Y	S	G	A	N	C	Q	Y		
S.purp.	61	C	A	S	A	P	C	C	L	N	G	G	S	C	I	D	G	V	N	G	Y	M	C	D	C	L	A	G	Y	T	G	T	H	C	E	T	
Pro.Z	1	C	A	S	Q	P	C	C	L	N	G	G	S	C	C	Q	D	S	I	R	G	Y	A	C	T	C	A	P	G	Y	E	G	K	N	C	E	L
Fact.X	61	C	E	G	H	P	C	C	L	N	Q	G	K	C	K	D	G	I	G	D	Y	T	C	T	C	L	E	G	F	E	G	K	N	C	E	L	
Fact.VII	61	C	E	T	S	P	C	C	L	N	Q	G	M	C	K	D	I	N	S	Y	E	C	W	C	Q	A	G	F	E	G	T	N	C	E	L		
Fact.IX	1	C	E	S	N	P	C	C	L	N	G	G	V	C	C	H	Q	H	R	E	S	F	S	C	D	C	P	P	Y	G	N	G	C	E	Q		
Lin-12	361	C	L	E	N	P	C	C	L	N	G	G	R	C	L	E	V	E	G	H	R	L	C	C	C	P	P	V	G	Y	T	G	P	F	C	D	V
Fact.XII	121	C	R	T	N	P	C	C	L	H	G	G	R	C	L	E	V	E	G	H	R	L	C	C	C	L	K	G	F	A	R	D	G	N	L	C	
Mu.egf	81	C	G	P	G	G	C	G	S	H	A	R	C	V	S	D	G	E	T	A	E	C	Q	C	C	C	L	K	G	F	A	R	D	G	N	L	C

69

[illegible]

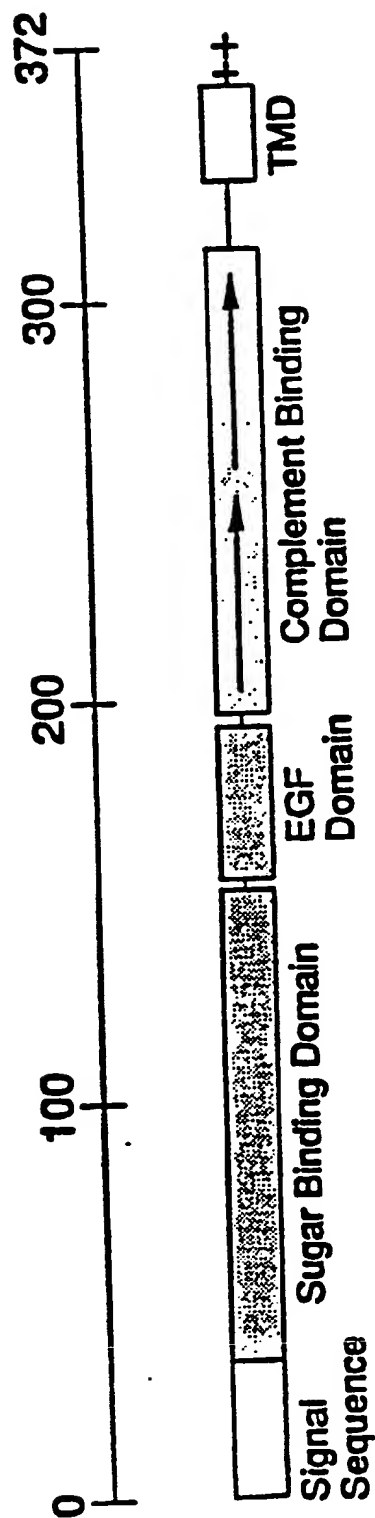
46	WS	S	P	E	P	I	C	Q	V	V	Q	C	E	P	L	E	A	P	E	L	G	..	I	M	D	C	I	M	P	..	L	G	N	F	S	..	F	O	
47	WT	N	D	I	P	I	C	E	V	V	K	C	L	P	V	T	A	P	E	N	G	K	I	V	S	S	A	M	E	P	..	D	R	E	Y	H	..	F	O
47	WI	N	D	I	P	L	C	E	V	V	K	C	L	P	V	T	E	L	E	N	G	R	I	V	S	Q	A	A	E	T	..	D	O	E	Y	..	F	O	
47	WS	P	E	L	P	V	C	A	P	I	C	P	P	S	I	P	T	F	A	..	T	L	R	V	Y	K	P	S	A	G	N	S	L	Y	R	..	L	O	
48	WS	P	A	A	P	T	C	E	V	K	S	C	D	D	F	M	Q	O	L	N	..	O	R	..	V	L	F	P	..	V	..	N	L	O	..	V	R		
48	WT	P	V	P	O	C	K	V	A	A	C	E	A	..	T	Q	R	O	L	..	L	T	..	K	P	..	L	T	..	K	P	..	O	H	O	..	F	..	T
48	WS	Q	T	E	P	I	C	R	O	P	Y	S	D	F	E	D	V	A	P	..	A	L	G	T	S	F	..	S	..	H	M	L	Q	A	..	T	N		
48	WS	Q	O	T	A	I	C	D	N	G	A	G	Y	C	..	S	N	P	G	..	P	I	G	T	R	K	V	..	Q	S	Q	Y	R	..	L	E			
48	VL	V	Q	E	M	H	G	C	K	I	A	C	..	V	L	P	V	L	M	..	G	I	Q	S	H	P	O	..	K	P	F	Y	..	V	O				
48	WS	N	S	K	L	K	C	O	P	I	V	D	C	G	P	E	S	I	E	N	G	..	K	V	E	..	D	P	..	E	S	T	L	..	E	G			
48	WR	P	S	P	T	C	E	K	I	T	C	R	K	P	D	V	S	H	Q	E	..	M	V	S	Q	F	O	P	..	I	..	Y	M	..	Y	K			
48	WS	T	A	V	E	F	C	K	K	S	C	P	..	N	P	O	E	I	R	..	N	Q	O	..	I	D	V	..	P	G	O	I	L	..	E	G			
48	WS	N	P	E	A	P	I	C	E	S	V	K	C	Q	S	P	P	S	I	S	N	G	..	R	H	N	G	Y	E	D	..	F	..	Y	T	..	D		

200 SKCA FNCSG
 210 QAVR FVCNSG
 220 QVVR FVCNSG
 230 DTAV FVCNSG
 240 AKVD FVCNSG
 250 PDVN SSGNSG
 260 PTQK TKE SSGNSG
 270 DSVT YHCSRGNSG
 280 EKVT VS CSRGNSG
 290 SVIR YTC EEPY YSGNSG
 300 DTIV FKCQKQNSG
 310 ATIS FSCNTQNSG
 320 SVVT YS CKN80

9/12

Fig. 7

PROTEIN MOTIFS IN THE LYMPHOCYTE HOMING RECEPTOR



10/12

Murine PLN Homing Receptor IgG Chimeras

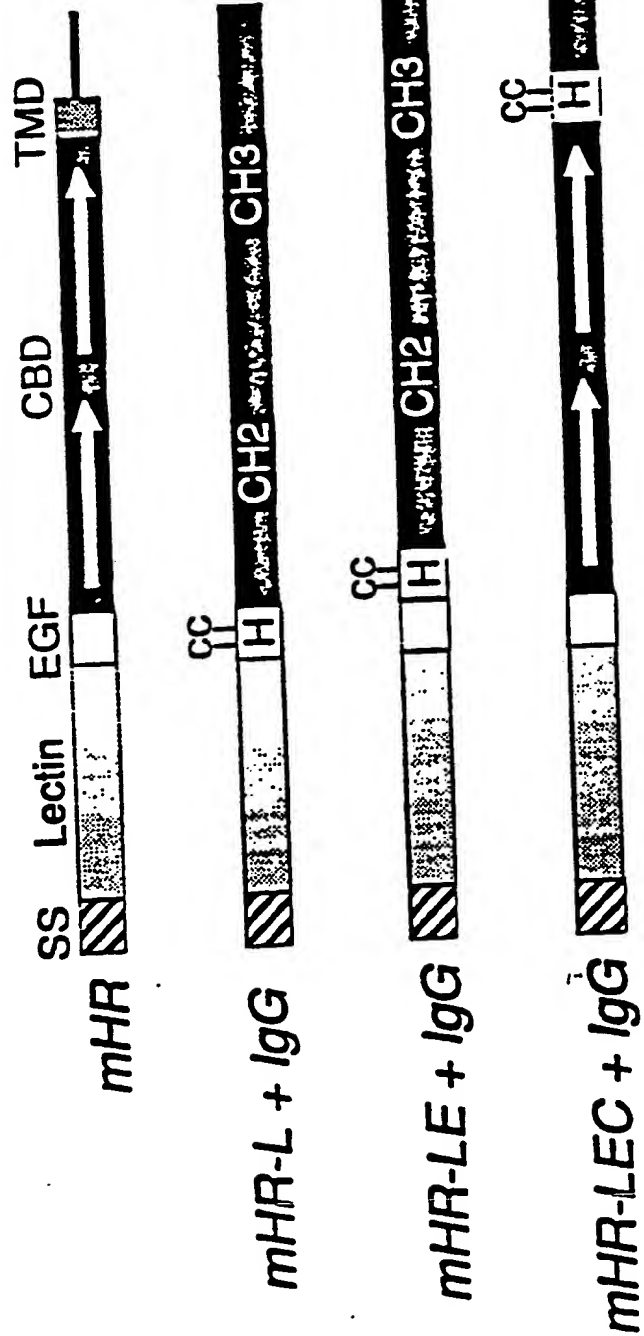


Fig. 8

11/12

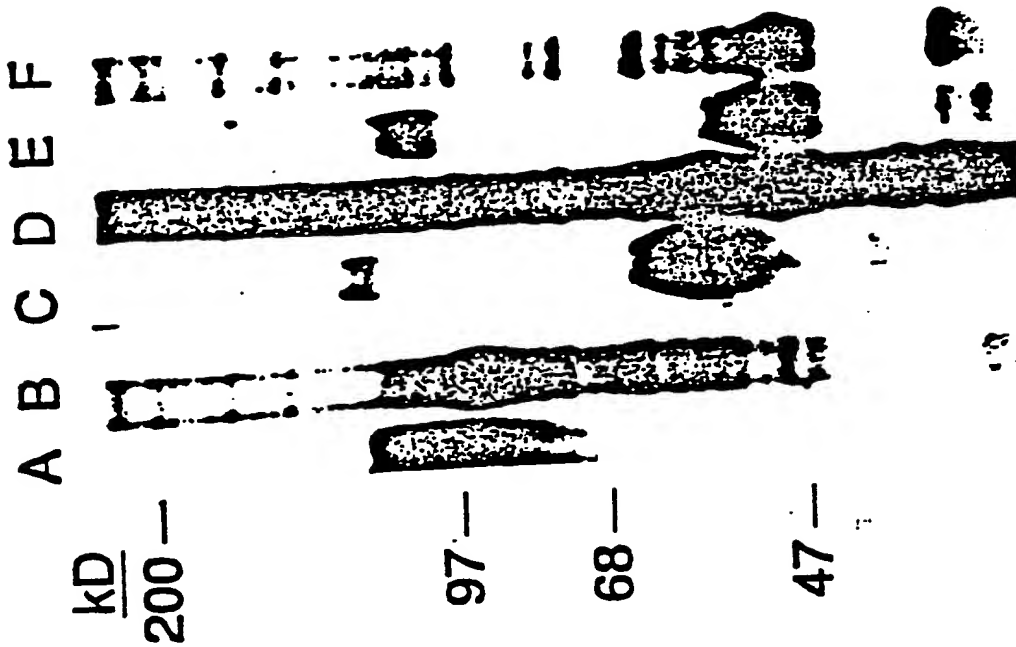
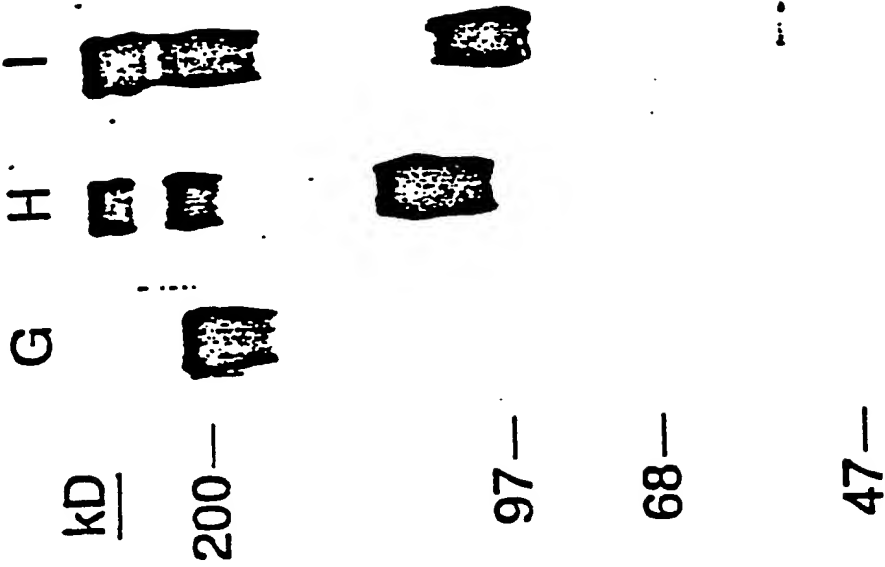
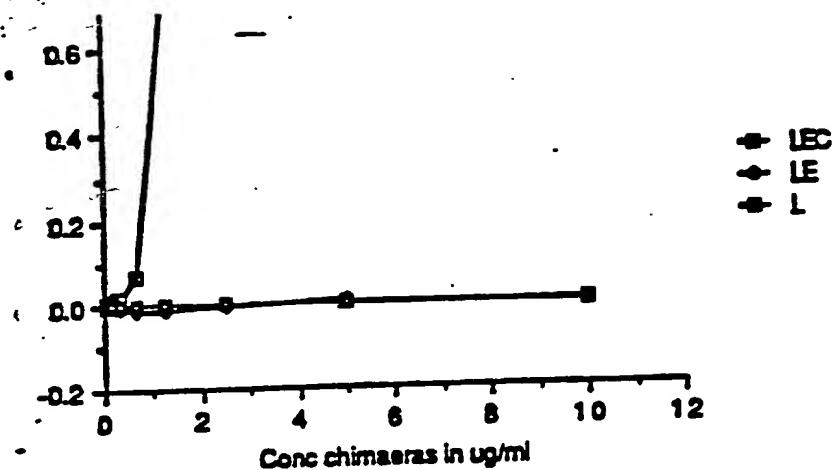


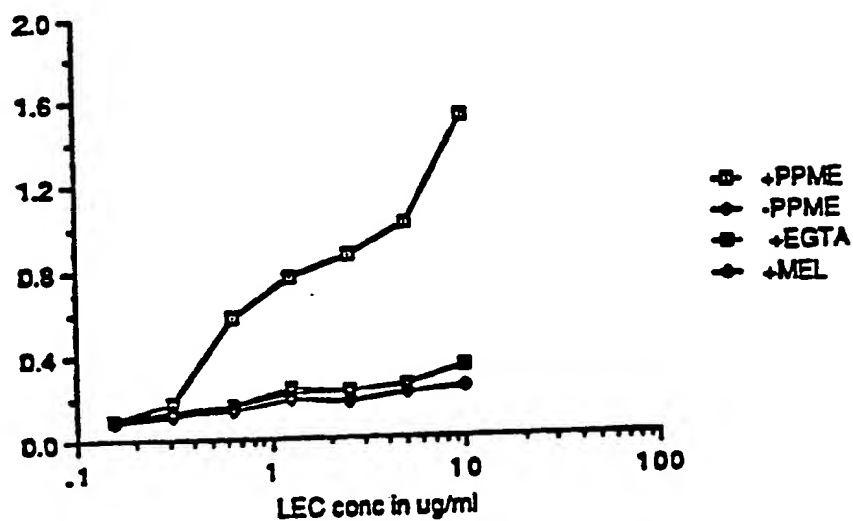
Fig. 9

12/12



Effects of EGTA and MEL-14 on PPME binding

Fig 10



Competition for PPME binding to LEC

